UNIVERSITY OF ALBERTA

Histone lysine methylation in mammalian development

by

Vincent L. Biron

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy in

Medical Sciences - Medical Genetics

Edmonton, Alberta
Spring 2007

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-29652-3
Our file Notre référence
ISBN: 978-0-494-29652-3

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

ABSTRACT

Histone lysine methylation has emerged as an important post-translational modification with the potential to index epigenetic information required for mammalian development. To investigate this potential, I examined the genome-wide dynamics of histone lysine methylation in mid-gestation mouse embryos. Using a panel of histone methyl-lysine antibodies, I determined the subnuclear distribution of 14 methyl-lysine derivatives of histones H3 and H4 *in situ*. This analysis revealed strikingly distinct genome wide distributions for some methyl-lysine derivatives, suggesting different functional properties. For instance, H3 K9 trimethylation (tMeK9-H3) exhibited dramatic cell cycle differences and appeared highly enriched in proliferative cells of the embryo. Likewise, H4 K20 monomethylation (mMeK20-H4) was also elevated in mitotic cells, but was more broadly distributed than tMeK9-H3. In contrast, the histone H4 K20 trimethyl derivative (tMeK20-H4) was progressively lost from dividing cells and became enriched in differentiating cells of the neural, and cardiac and skeletal muscle lineages, together with a concomitant loss of mMeK20-H4. These findings were corroborated in mouse primary cultures and C2C12 cells, which also revealed changes in histone H3 K36 and H3 K79 methylation during muscle differentiation. To further evaluate K20-H4 methylation changes during development, a mammalian retinogenesis model was used, enabling analysis of discrete neural cell lineages that arise from a single retinal progenitor cell population. Remarkably, progressive elevation of tMeK20-H4 was restricted to the ganglion cell layer and did not include other differentiated cell types, suggesting a lineage specific function.

To understand the mechanistic basis of histone H4-K20 methylation changes during myogenesis, published expression data and *in silico* gene expression analysis were used to

identify histone lysine methyltransferases that were enriched or restricted to the muscle cell lineage. Two of these enzymes, Prdm12 and Smyd1, were further examined by overexpression analysis in C2C12 differentiation and mouse primary limb bud cultures. Significantly, this revealed that Prdm12 appears to promote differentiation in addition to influencing the level of dMeK20-H4. Together, these results establish an unforeseen level of spatial and temporal differences in histone lysine methylation during development, and suggest that histone H4-K20 methylation is a key regulator of myogenic differentiation.

ACKNOWLEDGEMENTS

I wish to firstly thank my supervisor Dr. Alan Underhill for his guidance in several aspects of my graduate program, for teaching me to how to ask the right questions and for allowing me to pursue my own ideas. I would like to thank my committee members, Dr. Rachel Wevrick, for her constructive feedback, and Dr. Michael Hendzel for introducing me to nuclear structure and providing his expertise in this area. I would also like to thank Dr. Moira Glerum for her advice on my graduate program.

There is a long list of people I wish to thank for helping me with experiments. Thanks to Ning Hu and Gareth Corry who have helped me with virtually all aspects of my project, Dr. Kirk McManus, Dr. Sun and Darin Macdonald for their help with imaging, Dr. Fred Berry and Dr. Ramsey Saleem for providing experimental suggestions, and all others in the medical genetics department who have been helpful along the way.

In the early years of my graduate program, I met my wife Carolyn, who has been more helpful than she may realize in helping me complete this thesis. I thank her for her endless support and encouragement. I would also like to thank my daughter Elouise for giving me all the motivation I could ever need. I wish to thank my parents, my brothers Francois and Pierre and my sister Madelaine, whose memory helps me along the way.

Financial Support

I have been supported by the National Sciences and Engineering Research Council Post Graduate Scholarship B, Faculty of Medicine and Dentistry 75th Anniversary Award, Province of Alberta Graduate Student Award, Alberta Education Graduate Student Award, Walter H. Johns Graduate Fellowship, Mary Louise Imrie Graduate Student Award, and a Graduate Research Assistantship from the University of Alberta. My work has been supported by the Canadian Institutes of Health Research, Alberta Cancer Board and Alberta Heritage Foundation for Medical Research grants to DAU.

TABLE OF CONTENTS

CHAPTER 1 ♦ LITERATURE REVIEW **1**

INTRODUCTION 2

GENE REGULATION IN A DYNAMIC AND COMPARTMENTALIZED NUCLEAR ENVIRONMENT 3

THE NUCLEOSOME 4

THE HISTONE CODE WITHIN MULTIPLE LAYERS OF REGULATION 6

Histone post-translational modification effector domains *7*

Acetylation without a code *8*

Histone variants *10*

ENZYMATIC REACTIONS AND THE MACHINERY OF HISTONE METHYLATION 11

Histone arginine methylation *11*

Histone lysine methylation *12*

Histone lysine methylation in the DNA damage response *16*

Transcriptional initiation, elongation and termination *17*

Heterochromatinization *20*

LESSONS FROM HMTASE-DEFICIENT MICE 28

DEREGULATION OF HISTONE METHYLTRANSFERASES IN HUMAN DISEASE 34

Carcinogenesis *34*

Developmental disorders *39*

Cardiac disease *42*

OBJECTIVES AND RATIONALE 43

CHAPTER 2 ♦ ANIMALS, MATERIALS AND METHODS **51**

IMMUNOFLUORESCENCE 52

Mouse embryo immunofluorescence *52*

Chick embryo immunofluorescence *53*

Immunofluorescence with immortalized cultures *53*

Immunofluorescence with primary cultures *55*

Bromodeoxyuridine labeling of primary cultures 56

ANALYSIS USING KNOCKOUT MICE 56

Yolk sac DNA extraction and genotyping 57

IMAGE COLLECTION AND ANALYSIS 57

HISTONE ACID EXTRACTION AND IMMUNOBLOT ANALYSIS 59

IN SILICO ANALYSIS OF HISTONE METHYLTRANSFERASE GENE EXPRESSION 60

FIRST STRAND cDNA SYNTHESIS AND RT PCR 60

DNA CONSTRUCTS AND TRANSFECTIONS 61

Constructs 61

Transfections 62

CHAPTER 3 ♦ DISTINCT DISTRIBUTION AND DYNAMICS OF HISTONE METHYL-LYSINE DERIVATIVES IN MOUSE DEVELOPMENT 63

INTRODUCTION 64

RESULTS 65

Immunofluorescence analysis of histone lysine methylation in mouse embryogenesis 65

Histone H3 lysine-4 methylation 66

Histone H3 lysine-9 methylation 67

Histone H4 lysine-20 methylation 68

Regional differences in methyl histone derivatives 70

Spatiotemporal distributions of histone H4 lysine-20 methyl derivatives 71

K20-H4 methylation in caltreticulin deficient embryos 74

DISCUSSION 98

Dynamics of histone lysine methylation in development 100

CHAPTER 4 ♦ LINEAGE SPECIFIC GENOME-WIDE DISTRIBUTIONS AND DYNAMICS OF HISTONE LYSINE METHYLATIONS 105

INTRODUCTION 106

RESULTS 108

Histone lysine methylation dynamics during skeletal muscle differentiation 108

DISCUSSION 129

Distinct histone methyl-lysine residues are dynamic in skeletal muscle differentiation 129

CHAPTER 5 ♦ EXPRESSION AND DISTRIBUTION OF SMYD1 AND PRDM12 DURING MYOGENESIS 137

INTRODUCTION 138

RESULTS 140

In silico expression profiling of HKMTases 140

Prdm12 and Smyd1 in striated muscle development 142

Characterization of chromatin modifying enzymes in myogenesis 142

Expression dynamics of Smyd1 during skeletal muscle differentiation 142

DISCUSSION 170

Tissue specific roles for HMTases 170

HMTase regulation of myogenesis 171

Smyd1 in muscle development 172

A potential role for Prdm12 in myogenesis 176

CHAPTER 6 ♦ CONCLUSIONS AND FUTURE PERSPECTIVES 179

DISCUSSION 180

Histone lysine methylations demarcate distinct nuclear subdomains 180

Epigenetic control by tMeK20-H4 182

Smyd1 and Prdm12 in myogenesis 183

An epigenetic blueprint of mammalian embryogenesis 185

FUTURE PERSPECTIVES 189

HMTases and diet 189

Epigenetics of retinal development 189

Histone methylation in epigenetic disorders 190

Histone methyl-lysines as chemotherapeutic targets 192

APPENDICES 199

REFERENCES 202

LIST OF TABLES

Chapter 3

TABLE 3- 1: SUMMARY OF HISTONE POST-TRANSLATIONAL MODIFICATIONS EXAMINED IN MOUSE MID-GESTATION EMBRYO SECTIONS. 75

Chapter 5

TABLE 5- 1: SUMMARY OF *IN SILICO* GENE EXPRESSION ANALYSIS FROM 45 HMTASES. 151

TABLE 5- 2: IMMUNOFLUORESCENCE ANALYSIS OF FLAG-PRDM12 OVEREXPRESSION ON HISTONE LYSINE METHYLATIONS. 152

TABLE 5- 3: IMMUNOFLUORESCENCE ANALYSIS OF FLAG-PRDM12 OVEREXPRESSION ON MUSCLE CELL MARKERS. 153

Chapter 6

TABLE 6- 1: SUMMARY OF HISTONE LYSINE METHYLATION CHANGES OCCURRING DURING MOUSE EMBRYOGENESIS. 197

LIST OF FIGURES.

Chapter 1

FIGURE 1- 1: SUBNUCLEAR COMPARTMENTS 45

FIGURE 1- 2: HISTONE ACETYLATION, PHOSPHORYLATION AND METHYLATION. 47

FIGURE 1- 3: HISTONE MODIFICATION CHEMISTRY. 49

Chapter 3

FIGURE 3- 1: SUBNUCLEAR DISTRIBUTION OF HISTONE H3 MONO- AND TRI-METHYL LYSINE 4 IN MOUSE IMMORTALIZED, PRIMARY AND EMBRYO CELLS. 76

FIGURE 3- 2: SUBNUCLEAR DISTRIBUTION OF HISTONE H3 MONO- AND TRI-METHYL LYSINE 9 IN MOUSE IMMORTALIZED, PRIMARY AND EMBRYO CELLS. 78

FIGURE 3- 3: SUBNUCLEAR DISTRIBUTION OF HISTONE H4 MONO- AND TRI-METHYL LYSINE 20 IN MOUSE IMMORTALIZED, PRIMARY AND EMBRYO CELLS. 80

FIGURE 3- 4: HISTONE METHYL LYSINE DISTRIBUTION IN MOUSE NEUROEPITHELIUM. 82

FIGURE 3- 5: SPATIOTEMPORAL DYNAMICS OF H4 MONO- AND TRI-METHYL LYSINE 20 IN MOUSE NEURAL DEVELOPMENT. 84

FIGURE 3- 6: DYNAMICS OF MMEK20-H4 IN DEVELOPING CHICK RETINA. 86

FIGURE 3- 7: DYNAMICS OF TMEK20-H4 IN DEVELOPING CHICK RETINA. 88

FIGURE 3- 8: DIFFERENTIAL DISTRIBUTIONS OF MMEK20-H4 AND TMEK20-H4 IN MOUSE MYOTOME DEVELOPMENT. 90

FIGURE 3- 9: DIFFERENTIAL DISTRIBUTIONS OF H4 MONO- AND TRIMETHYL LYSINE 20 IN MOUSE CARDIAC MUSCLE DEVELOPMENT. 92

FIGURE 3- 10: DISRUPTED LEVELS OF HISTONE H4 TRIMETHYL LYSINE 20 IN CARDIAC MUSCLE OF CALRETICULIN NULL MICE. 94

FIGURE 3- 11: SUMMARY FOR HISTONE METHYLATION IN MOUSE EMBRYONIC DEVELOPMENT 96

Chapter 4

FIGURE 4- 1: IMMUNFLUORESCENCE OF SELECT HISTONE LYSINE METHYLATIONS SHOWING STATIC DISTRIBUTIONS THROUGHOUT C2C12 MUSCLE DIFFERENTIATION. 115

FIGURE 4- 2: IMMUNOFLUORESCENCE ANALYSIS OF MEK20-H4 OVER THE COURSE OF SKELETAL MUSCLE DIFFERENTIATION. 117

FIGURE 4- 3: IMMUNOFLUORESCENCE ANALYSIS OF MEK36-H3 OVER THE COURSE OF SKELETAL MUSCLE DIFFERENTIATION. 119

FIGURE 4- 4: IMMUNOFLUORESCENCE ANALYSIS OF MEK79-H3 OVER THE COURSE OF SKELETAL MUSCLE DIFFERENTIATION. 121

FIGURE 4- 5: IMMUNOBLOT ANALYSIS OF SELECT HISTONE METHYL-LYSINES IN DIFFERENTIATING C2C12 CELLS. 123

FIGURE 4- 6: MUSCLE SPECIFIC CHANGES IN H4 MONO- AND TRIMETHYL LYSINE 20 LEVELS IN MOUSE PRIMARY LIMB BUD CULTURES. 125

FIGURE 4- 7: HISTONE H4-K20 METHYLATION IN DIFFERENTIATING CHONDROCYTES DERIVED FROM MOUSE PRIMARY LIMB BUD CULTURES. 127

Chapter 5

FIGURE 5- 1: UNIGENE EXPRESSION PROFILE OF SMYD1 AND PRDM12 154

FIGURE 5- 2: SYMATLAS EXPRESSION PROFILE OF SMYD1 AND PRDM12. 156

FIGURE 5- 3: HMTASE EXPRESSION DURING C2C12 SKELETAL MUSCLE DIFFERENTIATION. 158

FIGURE 5- 4: SCHEMATIC OF PRDM12 AND SMYD1 DOMAIN STRUCTURE AND FUSION PROTEINS USED FOR THEIR FUNCTIONAL ANALYSIS. 160

FIGURE 5- 5: FUNCTIONAL ANALYSIS OF SMYD1 IN C2C12 CULTURES AND PRIMARY CELLS 162

FIGURE 5- 6: OVEREXPRESSION ANALYSIS OF HA-PRDM12 IN C2C12 CELLS. 164

FIGURE 5- 7: OVEREXPRESSION ANALYSIS OF FLAG-PRDM12 IN C2C12 CULTURES AND PRIMARY CELLS. 166

FIGURE 5- 8: LOCALIZATION OF FLAG-PRDM12 AND MARKERS OF MUSCLE DIFFERENTIATION. 168

Chapter 6

FIGURE 6- 1: FUNCTIONAL CHROMATIN DOMAINS DEFINED BY HISTONE LYSINE METHYLATION. 192

FIGURE 6- 2: SUBNUCLEAR RE-ORGANIZATION OF CHROMATIN THROUGH DYNAMIC K20-H4 METHYLATION DURING MUSCLE DIFFERENTIATION. 195

Appendices

FIGURE A-1: SUBNUCLEAR DISTRIBUTION OF HISTONE TMEK4-H3 IN COS-7 CELLS AND MOUSE EMBRYO SECTIONS. 199

FIGURE A-2: CELL CYCLE ANALYSIS OF K20-H4 METHYLATIONS IN DIFFERENTIATING MOUSE PRIMARY LIMB BUD CULTURES. 201

LIST OF ABBREVIATIONS, SYMBOLS AND NOMENCLATURE

Ac	acetyl
ADP	adenosyl diphosphate
AML	acute myeloid leukemia
Ash1	absent, small or homeotic imaginal disks 1
ATP	adenosine triphosphate
Brd	bromodomain
BrdU	bromodeoxyuridine
CBP	CREB binding protein
ChIP	chromatin immunoprecipitation
CML	chronic myeloid leukemia
DAPI	4',6'- diamidino phenylindole
DIC	differential interference contrast
DNA	deoxyribonucleic acid
ES	embryonic stem
FBS	fetal bovine serum
FITC	fluorescent isothiocyanate
GFP	green fluorescent protein
HA	hemagglutinin
HAT	histone acetyltransferase
HDAC	histone deacetylase

HKDT	histone lysine demethylase
HKMT	histone lysine methyltransferase
HMT	histone methyltransferase
HP1	heterochromatin protein 1
LSD1	lysine specific demethylase 1
Me	methyl
MHC	myosin heavy chain
MLL	mixed lineage leukemia
NSD	nuclear set domain
PBS	phosphate buffered saline
PcG	polycomb group
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
Ph	phosphoryl
PHD	plant homeodomain
PML	promyelocytic leukemia
PR	PRDI-BF1 and RIZ
PRC	polycomb repressive complex
PRMT	protein arginine methyltransferase
PWWP	proline-tryptophan-tryptophan-proline
RNA	ribonucleic acid
SDS	sodium dodecyl sulfate

SET	Su(var), Enhancer-of-zeste and Trithorax
SMYD	set and mynd domain
TBE	tris–boric acid–EDTA
TBS	tris-boric acid-saline
TRITC	tetrarhodamine isothiocyanate
Ub	ubiquitin
WHS	Wolf-Hirschhorn syndrome
YFP	yellow fluorescent protein
Znf	zinc finger

Chapter 1 ♦ Literature review

Introduction

Eukaryotic gene regulation occurs within an extremely complex and plastic nuclear environment. For instance, DNA is compacted by proteins in a chromatin fiber that can undergo large scale rearrangements to alter the expression of genes. In addition, chromatin interacts with several subnuclear domains that may serve to spatially coordinate enzymatic reactions and consequently enhance their efficiency. As the structure and function of the nucleus continues to be elucidated, a more physiological understanding of gene regulation is emerging.

It is well appreciated that epigenetic information, which is not encoded in genomic DNA sequence, plays a significant role in gene regulation. Epigenomic information is embedded within chromatin, and includes DNA methylation, RNA, and a gamut of histone post-translational modifications. Because the epigenome has powerful control over gene expression, enormous efforts have been directed toward understanding its various mechanisms of regulation. In recent years, histone lysine methylation in particular has become a major focus of study with these modifications having been shown to play fundamental roles in development and cancer. This thesis examines the role of various histone methylation states in mammalian development.

The dissection of the epigenome is providing novel insight into mechanisms of development and disease. Importantly, the epigenome offers an accessible protein platform for chemotherapeutic agents that may serve to inhibit or enhance deregulated genes. Further characterization of epigenetic mechanisms will undoubtedly advance modern medicine.

Gene regulation in a dynamic and compartmentalized nuclear environment

Over one meter of DNA is packaged as chromatin within a nuclear diameter of ~ 10 microns (Horn and Peterson 2002). Chromatin fibers ranging from 30-300 nm in diameter fill the majority of the nuclear volume. Thicker chromatin fibers that generally do not undergo transcription are commonly found at the nuclear and nucleolar periphery, whereas transcriptionally active chromatin is dispersed throughout the nuclear interior (Carmo-Fonseca 2002). Interphase chromosomes are also organized within distinct 'chromosome territories' (Cremer and Cremer 2001). Within this steady-state organization, chromatin is highly dynamic and can undergo rapid changes in compaction. In addition, gene loci have been shown to travel between nuclear microenvironments that have distinct gene regulatory properties (Lamond and Earnshaw 1998; Belmont 2003).

The interchromatin space is populated by well-defined subnuclear compartments that orchestrate various functions (Spector 2001). For instance, nucleoli are sites of ribosome biogenesis and are intimately associated with chromatin through ribosomal gene loci (Carmo-Fonseca 2002). Splicing speckles are large domains with an elevated concentration of pre-mRNA splicing factors (Lamond and Spector 2003). The splicing speckle periphery is enriched in active chromatin, an organization that conceivably provides more efficient co-transcriptional splicing of multi-exon genes. PML domains have been implicated in regulating various activities ranging from transcription to DNA repair, and it appears that chromatin contacts are also made by these domains (Ching et al. 2005). Live cell imaging studies that have traced the movements of PML proteins have

illustrated that nuclear domains may have distinct mobility according to their position within the nucleus, and may also be responsive to toxic metabolic states (Eskiw et al. 2004; Gorisch et al. 2004). Such nuclear domain dynamics are regulated in part by chromatin, which can act locally to constrain nuclear proteins through direct protein-chromatin attachments or by forming local enclosures that restrict the diffusion of proteins. Taken together, these examples demonstrate that chromatin-dependent processes occur within a highly regulated and dynamic nuclear environment.

Of particular importance to this thesis, it should be noted that nuclear organization varies between different cell types and as such only basic principles can be applied to mammalian cells with respect to nuclear structure. By electron microscopy, significant differences can be seen in the organization of chromatin between terminally differentiated nuclei of different cell types (Francastel et al. 1999). In addition to these differences, chromosome territories occupy non-random positions that are cell-type specific. Though it is still debated how these large-scale chromatin differences are established (Taddei et al. 2004), further characterization of chromatin-dependent processes will contribute to our understanding of mammalian development.

The nucleosome

The nucleosomal model of chromatin structure originated approximately 30 years ago, after much opposition from supporters of the super-coil model, whereby the DNA helix is constrained to form a superhelical fiber (Pardon and Wilkins 1972). In 1973, an anonymous review of a manuscript by D.F.L. Woodcock stated the following:

A eukaryotic chromosome made out of self-assembling 70A units, which could perhaps be made to crystallize, would necessitate rewriting our basic textbooks on cytology and genetics! I have never read such a naïve paper purporting to be of such fundamental significance. Definitely it should not be published anywhere! (Van Holde 1989)

The nucleosome is now known to form the base unit for chromatin assembly, comprised of an H3:H4 tetramer and two H2A:H2B dimers in an octameric complex that is encircled almost twice by approximately 146 base pairs of DNA (Luger et al. 1997). The amino and carboxyl-termini of histones are often shown as free-floating tails emanating from the nucleosome, as portrayed by the crystal structure model of Luger et al. It should be noted, however, that this structure does not include linker DNA and is derived from histones lacking post-translational modifications. Nevertheless, the nucleosomal structure appears to be highly conserved and is repeated every 200 +/- 40 bp in all eukaryotic genomes (Luger et al. 1997). In addition, histone proteins appear to be conserved as far as archaebacteria, which have histone proteins that form nucleosome-like structures (Pereira et al. 1997; Pereira and Reeve 1998).

Nucleosomes assemble into higher order chromatin structures that are stabilized by the binding of histone H1 - the so-called 'linker' histone - where DNA enters and exits the nucleosome. The recent crystallization of a tetranucleosome suggests that nucleosomes are compacted according to the crossed-linker model, whereby "linker DNA zigzags back and forth between two stacks of nucleosome cores" (Schalch et al. 2005). In the interphase nucleus, the majority of chromatin fibers appear to be greater than 30 nm in diameter - a level of compaction that is believed to make DNA inherently inaccessible to regulatory

proteins (Horn and Peterson 2002). The mechanisms by which gene regulatory factors contend with such a repressive chromatin environment remain a complex biological problem that is beginning to be unraveled. Because gene promoter elements *in vivo* are not generally found in an accessible, linear 'beads-on-a-string" conformation, chromatin fibers would need to be locally unraveled in order to be recognized by transcription factor DNA-binding domains. One solution to this problem is for transcription factors to act in partnership with chromatin regulatory proteins that may act to either decompress or compress chromatin to either activate or repress genes, respectively. Alternatively, some DNA-binding transcription factors could access specific promoter elements by carrying intrinsic chromatin remodeling domains.

The histone code within multiple layers of regulation

In recent years, the nucleosome has emerged as an active participant in gene regulation by providing epigenetic information through a range of post-translational modifications found within histone globular domains and their tails (Cosgrove and Wolberger 2005). These include acetylation, methylation, phosphorylation, ubiquitination, and ADP-ribosylation, which are classified under a recently adopted nomenclature (Berger 2002; Yamane et al. 2006). Occurring most commonly on amino-terminal histone tails, these modifications are thought to provide a combinatorial 'histone code' that can be read by nuclear proteins to influence activities such as transcription, replication and cell cycle regulation (Strahl and Allis 2000; Turner 2000Strahl, 2000 #5; Jenuwein and Allis 2001; Turner 2002). Based on the range of modifications and the number of amino acid residues in the four core histones that are now known to be substrates, the potential exists to convey an enormous volume of

epigenetic information (Fischle et al. 2003). When also taking into account the effect of neighboring nucleosomes, histone variants and histone tail proteolysis, the hypothesis that histones provide a combinatorial readout for biological processes becomes extremely complex. In addition, acetylation does not appear to provide a combinatorial readout, but functions cumulatively to influence chromatin structure (Henikoff 2005). Thus, although chromatin modifications are important effectors of distinct biological processes, they should perhaps not always be viewed as a bar code that can be read in a combinatorial fashion. Nevertheless, histone post-translational modifications can serve as docking sites for regulatory proteins though modification-specific domains, as will be described below.

Histone post-translational modification effector domains

Evidence for the existence of a histone code comes from several examples demonstrating that histone post-translational modifications can be read by specific binding modules, exemplified by the bromodomain and chromodomain (Yang 2005). Bromodomain proteins can specifically recognize acetyl-lysine containing motifs found in both histone and non-histone proteins, and can mediate several chromatin-dependent processes (Yang 2004). Some histone acetyltransferases (HATs), namely CBP, p300, PCAF, and GCN5 catalyze the addition of single acetyl groups onto histone lysines, and also dock onto histone acetylated residues. Brd proteins contain double bromodomains that can simultaneously recognize two acetyl-lysines, which serve to remodel distinct chromatin domains (Yang 2004). ATP-dependent chromatin remodeling and nucleosome assembly proteins also direct specific functions via bromodomain recognition of distinct histone acetylations. In addition, bromodomains can be found in histone lysine methyltransferase

(HKMTase) enzymes such as Ash1, PR-domain members, and the MLL subfamily (de la Cruz et al. 2005). Chromodomain and Chromo-related domains can recognize specific histone lysine methylation residues and also appear to interact with RNA and DNA (Brehm et al. 2004). Heterochromatin Protein 1 (HP1) is the prototypical chromodomain protein, which specifically recognizes K9-H3 methylations to induce and stabilize the formation of higher-order chromatin structures that are transcriptionally silent. The extensive characterization of HP1 has lead to the identification of chromodomain residues that are of predictive value with respect to methyl-lysine binding, but the structural basis of DNA and RNA binding motifs remains to be characterized (Brehm et al. 2004). Although chromodomain and chromo-related domains appear more restrictive than bromodomains in their biological functions, they can also be found in HKMTases, HATs, and chromatin-remodeling proteins (de la Cruz et al. 2005). The examples above provide mechanistic support for the histone code hypothesis, but some aspects of histone biology are inconsistent with the originally proposed hypothesis and are outlined in the next section.

Acetylation without a code

Histone acetylation is currently the best characterized histone modification, and has been shown to influence various chromatin-templated processes including transcription, DNA replication and DNA repair. Several conserved sites of acetylation can be found on the tails of histones H3 (K9, K14, K18, K23 and K27) and H4 (K5, K8, K12 and K16) along with less conserved residues on histones H2A and H2B (Kurdistani et al. 2004). Histone acetylation-dependent processes can be thought of categorically as functioning on two levels. The first occurs by influencing the folding properties of the chromatin fiber, and the

second by providing recruitment platforms for chromatin regulatory proteins as discussed above. In general, hyperacetylation of chromatin is thought to generate open structures that are conducive to transcriptional activity whereas hypoacetylation is associated with heterochromatic and transcriptionally inactive loci. Various histone acetylation patterns are established by the actions of histone acetyltransferases (HATs) and histone deacetylases (HDACs), which generally operate as transcriptional activators and repressors, respectively. Both HATs and HDACs comprise large protein families that work in combination with other regulatory proteins that influence a broad spectrum of activities, which have been extensively characterized (Verdone et al. 2005) but are beyond the scope of this review.

Histone acetylation does not appear to have properties that are consistent with being part of a 'histone code' (Henikoff 2005). The histone code hypothesis predicts that "multiple histone modifications, acting in a combinatorial or sequential fashion on one or multiple histone tails, specify unique downstream functions" (Strahl and Allis 2000). To date, the influence of histone acetylation sites on transcription has neither been combinatorial nor has it been consistent in different gene contexts (Kurdistani et al. 2004). As recently demonstrated by Dion et al., histone acetylation appears to act cumulatively (Dion et al. 2005). By mutating acetylation sites on H4, this study found that the position of the mutation was not as important as the number of mutations for determining gene activity. However, acetylation of K16-H4 appeared to behave somewhat autonomously, consistent with previous studies demonstrating a specific role for this acetylation site such as X-chromosome dosage compensation in Drosophila, and the generation of specific

binding sites for Sir3 or Bdf1 transcription factors in yeast (Kurdistani et al. 2004; Dion et al. 2005; Henikoff 2005). Taken as a whole, it is evident that a 'histone code' cannot be generalized to all histone modifications, although certain modifications may show a consistent association with a particular function.

Histone variants

Canonical core histones H3 and H2A have variant isoforms that are associated with distinct chromatin states (Henikoff et al. 2004). Histone H3.3 differs from H3 by only four amino acids within its globular domain and is present in transcriptionally active chromatin domains (Henikoff et al. 2004; Mito et al. 2005; Sarma and Reinberg 2005). This H3 variant is deposited during transcription in a replication-independent manner with a protein complex involving the HIR histone cell cycle regulation defective homolog A (HIRA) (Tagami et al. 2004). CENP-A is a centromeric H3 (CenH3) variant that is required for kinetochore assembly, and it is consistently associated with centromeres throughout eukaryotes, making it a defining feature of this structure (Henikoff et al. 2004; Sarma and Reinberg 2005). The amino terminus of CENP-A differs considerably from that of H3 and this variant also contains a unique centromeric recognition domain (Sarma and Reinberg 2005). Interestingly, the *S. pombe* CenH3 variant Chromosome Segregation 4 (Cse4) is restricted to the centromere through a mechanism that involves ubiquitin-mediated degradation of mislocalized CenH3 (Collins et al. 2004). In terms of histone H2, the H2AZ variant is elevated in euchromatin, where it is thought to destabilize the nucleosome to make it more amenable to transcriptional activation (Henikoff et al. 2004). The inactive X chromosome is characterized by an elevation of the H2A variant macro H2A and a

depletion of H2A Barr body-deficient (H2A Bbd), an H2A variant that is widely dispersed with a potential role in transcriptional activation (Henikoff et al. 2004; Sarma and Reinberg 2005). Together, these histone variants add an additional layer of complexity to gene regulation. It should be noted that an H4 variant has not been identified and the study of epigenetic processes involving this histone may therefore be more straightforward.

Enzymatic reactions and the machinery of histone methylation

In terms of conveying long-term epigenetic information, histone post-translational modifications can be limited by a relatively short half-life. For instance, histone acetylation generally has a turnover rate in the order of minutes (Waterborg 2002). In relation to other histone post-translational modifications, histone methylation is thought to be extremely stable with the potential to convey long-term epigenetic information. The recent discovery of methyltransferase enzymes that specifically catalyze the addition of methylations onto histone residues has provided a way to mechanistically understand histone methylation and its biological significance. These histone methyltransferases (HMTases) covalently add methyl-groups onto histone arginine and lysine residues in a pathway that is dependent on S-adenosylmethionine (SAM), which is ultimately derived from dietary methyl-donors (Huang 2002). Specific examples of these enzymes and their functional roles are provided in the following sections.

Histone arginine methylation

Histone arginines are methylated by protein arginine methyltransferases (PRMTs) that can catalyze the formation of monomethyl or dimethyl arginines in either symmetric or

asymmetric conformations (Cheng et al. 2005). In mammals, at least seven PRMT genes have been identified, some of which appear to methylate non-histone substrates. For instance, in addition to histones H4 and H2B, PRMT1 can methylate RNA-processing proteins and various other transcription factors (Cheng et al. 2005).

From a functional standpoint, histone arginine methylation has been shown to be involved primarily in transcriptional activation. For example, PRMT1 and PRMT4/CARM1 can methylate arginines of H4 and H3, respectively and cooperate with histone acetyltransferases to transcriptionally activate hormone-responsive promoters (Cuthbert et al. 2004; Cheng et al. 2005). This effect was recently shown to be antagonized by human peptidylarginine deiminase 4 (PAD4) (Cuthbert et al. 2004; Wang et al. 2004). PAD4 can convert either unmodified or monomethyl (but not dimethyl) arginine to citrulline, thereby inhibiting transcriptional activation. In this fashion, monomethyl arginine could represent a transcriptionally poised state which could be prevented from initiation by PAD4 deimination or become resistant to PAD4 by dimethylation (Denman 2005). Although this mechanism does not represent the action of a true demethylase, the final outcome results in the loss of arginine and the gain of a new histone modification. It therefore remains possible that citrulline conveys a specific function that remains to be identified, adding to the ever expanding complexity of histone biology.

Histone lysine methylation

Histone lysine methylation has important functions in chromatin-dependent processes such as transcription, DNA replication, and DNA repair. Histone N-terminal tails can be methylated on H3 at K4, K9, K27 and K36 and at a single site on H4 at K20. In addition,

H3 can be methylated in its core at K79. Methylation of one other histone lysine residues has been reported, but its significance is not known (Cocklin and Wang 2003). Each lysine residue can occur in mono-, di- or trimethylated states and appears to associate with distinct chromatin subsets. These methylations are added by histone lysine methyltransferases (HKMTases) that may have specificity for single or multiple lysines and/or methylation states (Table 1-1). As will be discussed, the dogma of histone methyl-lysine biology was revised following the identification of a histone lysine demethylases (Shi et al. 2005; Yamane et al. 2006).

With the exception of Dot1, all HKMTases contain a conserved PR or SET domain, which share ~ 20-30% amino acid identity (Huang 2002). The SET domain was originally identified as 120-150 amino acids of sequence homology shared between three *Drosophila* proteins involved in chromatin regulation, namely Su(var)3-9, Enhancer of Zeste and Trithorax (Cheng et al. 2005). The PR domain was identified by homology shared between the PRDI-BF1 and RIZ1 proteins, and was later shown to be similar to the SET domain (Huang et al. 1998). Interestingly, the SET domain is present in genomes from yeast to humans, and also in some prokaryotes, whereas the PR domain is not found in lower organisms. This suggests the PR domain may have evolved from the SET domain and may serve more specialized functions that are relevant to metazoans (Huang 2002).

Structural characterization of SET domain proteins has provided some insight of their function. Four conserved motifs in the SET domain appear to be responsible for the three steps in histone methyl-lysine catalysis, namely SAM binding, target lysine binding and methyl transfer (Cheng et al. 2005). The SET domain active site forms a knot-like

structure and a 'tyrosine/phenylalanine switch' appears to control the product specificity for the number of methyl groups added (mono-, di or tri) (Cheng et al. 2005; Yamane et al. 2006). It is interesting to note that most PR domain proteins have a valine at the phe/tyr switch position which may be of functional significance. Although structural information is not available for PR domain proteins, an analysis of PR domain-containing 2 (PRDM2/RIZ1) unstructured segments suggests PR and SET domains have some structural homology (Derunes et al. 2005).

Until recently, it was unclear whether the enzymatic removal of histone methyl-lysines could occur because a histone lysine demethylase had not been identified. Although the presence of histone demethylases had been speculated for ~ 40 years, a true histone demethylase remained elusive until Shi et al. revealed the function of LSD1 (Shi et al. 2005). LSD1, transcriptional corepressor that is conserved from yeast to humans, was first found to specifically demethylate m/dMeK4-H3, but not tMeK4-H3, via an oxidation reaction. Three recent papers, by Shi et al., Lee et al. and Metzger et al. have now expanded the understanding of LSD1 function (Wysocka et al. 2005). Shi et al. have addressed the m/dMeK4-H3 demethylase activity of LSD1, revealing that its function is dependent on several other regulatory proteins such as HDAC1/2 histone deacetylases, the CoREST co-repressor and the BHC80 PHD-domain protein (Shi et al. 2005). In addition, this study showed that CoREST/LSD1 mediated demethylation is enhanced by hypoacetylation, suggesting synergy between histone deactylation and demethylation. Lee et al. further support the important role of regulatory proteins such as CoREST in directing the activity of LSD1 and its demethylase functions (Lee et al. 2005). Importantly, Metzger

et al. extend the function of LSD1 by demonstrating that it can function to demethylate m/dMeK9-H3 to de-repress androgen receptor target genes (Metzger et al. 2005). More recently, a separate histone demethylase termed JmjC domain-containing histone demethylase 1 (JHDM1) has been identified, which uses a different demethylation mechanism to remove m/dMeK36-H3 (Yamane et al. 2006). In addition, a related protein named Jmjd2b appears to mediate the removal of tMeK9-H3 (Fodor et al. 2006). Thus, both active and repressive histone lysine methylation marks can be removed by demethylase activity, and these recent findings will likely propel the rapidly advancing field of histone methyl-lysine biology.

In addition to histone demethylases, histone lysine methylation can also be lost by other mechanisms. One possibility is that histone H3 may undergo proteolytic cleavage, as shown in Tetrahymena H3 N-terminal tails and in the regulation of yeast centromeric H3 (Allis et al. 1980; Collins et al. 2004). Replacement of histones with histone variants could also be an effective mechanism to remove stable methylation marks (Henikoff et al. 2004). Such a mechanism is thought to operate in transcriptionally active chromatin whereby histone H3.1 is actively replaced by H3.3. It should be noted that to date, no histone variant has been identified for histone H4 and therefore modifications on this histone tail may not be subject to variant replacement.

Another important finding in our understanding of histone methyl-lysine biology was the discovery that, similar to PRMTases, many HKMTases have non-histone substrates. The Set7/9 HKMTase was recently shown to methylate lysines in p53 and TAF10, in addition to K4 on histone H3 (Chuikov et al. 2004; Kouskouti et al. 2004). In fact, Set7/9

can methylate p53 with higher affinity than H3-K4, which increases p53 stability (Chuikov et al. 2004). Similarly, monomethylation of TAF10 by Set7/9 increases the affinity of TAF10 for RNA polymerase II (Kouskouti et al. 2004). In addition, HKMTases may have non-histone cytoplasmic targets, as demonstrated by the Enhancer of zeste homolog 2 (Ezh2) protein. In T-lymphocytes, Ezh2 functions in a cytoplasmic HKMTase complex that acts on Vav1, a Rho GDP-GTP exchange factor that acts downstream of the T-cell receptor and promotes actin reorganization (Nolz et al. 2005; Su et al. 2005). In fibroblasts, Ezh2 also has an HKMTase-dependent cytoplasmic signal transduction function that acts on Vav2/3, which regulates actin. Taken together, it should be emphasized that histone lysine methylation can act beyond histone-dependent processes, and is likely to have several unidentified nuclear and cytoplasmic functions.

Histone lysine methylation in the DNA damage response

An accumulating body of evidence suggests histone lysine methylation has an important role in the DNA damage response by recruiting proteins to double-stranded DNA breaks (DSBs). Upon the induction of DSBs, different protein classes are recruited to the breakage site and can initiate damage checkpoint signaling (Vidanes et al. 2005). One such class is a set of adaptor proteins represented by the *S. cerevisiae* Rad9, *S. pombe* Crb2 and human 53BP1 checkpoint kinases, which share similar functions but are not highly homologous (Vidanes et al. 2005). In mammals, 53BP1 directly binds to dMeK79-H3 at DSB foci, an association that is lost in cells deficient for the K79-H3 HKMTase Dot1 (Huyen et al. 2004). The 53BP1 protein contains two Tudor domains, which are similar to chromodomains found on other methyl-binding proteins such as HP1. These Tudor

domains could therefore serve to target 53BP1 to MeK79-H3 (Charier et al. 2004; Sanders et al. 2004). Similarly, MeK79-H3 appears to serve as a binding site for Rad9 in *S. cerevisiae* (Vidanes et al. 2005). In *S. pombe*, recruitment of Crb2 to DSBs is dependent on MeK20-H4, which is regulated by the HKMTase Set9 (Sanders et al. 2004). Fission yeast Set9 mutants have a depletion of MeK20-H4 and a loss of cell-cycle checkpoint control. Unexpectedly, *S. pombe* H4-K20 methylation does not appear to participate in centromere structure or gene repression as it does in metazoans. With regards to MeK20-H4 function, this finding is intriguing because fission yeast and metazoans share several other molecular features of centromere structure and gene silencing. This could suggest that MeK20-H4 in metazoans serves a more specialized function that is beyond its role in heterochromatin, independent of MeK9-H3.

Transcriptional initiation, elongation and termination

Gene transcription is positively correlated with histone methylations involving K4, K36 and K79 on histone H3. Of these methylations, the role of MeK4-H3 in transcriptional activation has been more clearly established and widely appreciated. Chromatin immunoprecipitation (ChIP) studies have demonstrated that tMeK4-H3 is associated exclusively with transcriptionally active loci in both yeast and mammalian cells, and its levels peak at the promoters and early coding regions of active genes (Santos-Rosa et al. 2002; Schneider et al. 2004; Bannister et al. 2005). In contrast, mMeK4-H3 can be found in more downstream regions throughout the open reading frame (Kouskouti and Talianidis 2005; Morillon et al. 2005). Enzymatically, d/tMeK4-H3 is formed by Set1, which is recruited to newly initiated RNA polymerase II (RNAPII) transcription complex. The

degree of Set1 methylation is regulated by its association with an octameric protein complex known as Complex Proteins Associated with Set1 (COMPASS) (Krogan et al. 2003; Morillon et al. 2005; Yamane et al. 2006). In addition, COMPASS-mediated Set1 methylation is dependent on ubK123-H2B, which could either act as a 'wedge' in chromatin to promote Set1 access or as a 'bridge' that functions as a docking site for the enzyme (Henry and Berger 2002). The presence of Set1-directed histone modifications also facilitates binding of the chromatin remodeling ATPase Isw1p, leading to changes in chromatin structure that are required for transcription and recruitment of the cleavage and polyadenylation factor Rna15p (Santos-Rosa et al. 2003). In this regard, there is evidence for the association of both RNA 3' processing proteins and active chromatin with the splicing speckle periphery in mammalian cells (Schul et al. 1998). Moreover, an association of active chromatin with the splicing speckle periphery has also been demonstrated. Hyperacetylated histone H3, transcriptionally active RNA polymerase II and nascent transcripts are enriched adjacent to the splicing speckle, with splicing-associated proteins cycling between the speckle and sites of transcription (Hendzel et al. 1998; Szentirmay and Sawadogo 2000; Lamond and Spector 2003). Consistently, we have observed a specific enrichment of tMeK4-H3 at the splicing speckle periphery, suggesting a functional association between this methylation and the RNA splicing machinery (see appendix).

In yeast, the K36-H3 methyltransferase Set2 has been shown to associate with elongating RNAPII, suggesting that MeK36-H3 is important for elongation (Kizer et al. 2005). Set2 can also associate with COMPASS to coordinate early post-initiation events

with Set1 through d/tMeK36-H3 and tMeK4-H3 marks (Morillon et al. 2005). In metazoans, however, d/tMeK36-H3 levels are highest at the 3'end of genes, suggesting a role for this modification in transcriptional termination and/or early RNA processing (Bannister et al. 2005). Consistent with this hypothesis, Set2 contains a WW motif shared among other pre-mRNA splicing proteins (Lin et al. 2004). Because Set2 is also known to interact with RNAPII (Bird et al. 2004), it is tempting to speculate that this HKMTase could function to coordinate co-transcriptional splicing via MeK36-H3.

Lysine 79 is positioned within the histone H3 globular domain and is methylated by the non-SET/PR domain HKMTase Dot1 (Feng et al. 2002; van Leeuwen et al. 2002; Yamane et al. 2006). K79-H3 is hypomethylated in yeast telomeric and silent mating-type loci, and is depleted from heterochromatin in mammals (Yamane et al. 2006). Moreover, K79-H3 hypomethylated yeast chromatin recruits Sir proteins, which appear to stabilize this heterochromatic profile by preventing the deposition of MeK79-H3 (Yamane et al. 2006). In contrast to tMeK4-H3, the highest levels of dMeK79-H3 are found in 3' regions of mammalian genes and Dot1-directed deposition of this mark is also dependent on ubK123-H2B (Shilatifard 2004; Kouskouti and Talianidis 2005).

Histone methyl-lysines associated with transcription may provide a memory of active chromatin domains. In yeast, tMeK4-H3 has been shown to persist for several hours at gene promoters but is lost over a longer time scale, and has therefore been suggested to provide a form of short-term memory for transcription (Yamane et al. 2006). Distinct H3 K4, K36 and K79 methylation profiles have also been speculated to provide RNAPII with a transcriptional memory that could provide regulation of its activity (Shilatifard 2004). In

mammalian cells, tMeK4-H3 and dMeK79-H3 modifications have been shown to persist over the course of mitosis following transcriptional inhibition (Kouskouti and Talianidis 2005). Importantly, the stability of these modifications appears to be significantly longer than in yeast and is on a time scale that could permit the transmission of these methylations to daughter cells in developing organisms.

Heterochromatinization

The functional significance of heterochromatin in epigenetic regulation is often exemplified by position effect variegation (PEV) (Wakimoto 1998). This phenomenon was first described in studies of the *Drosophila* white eyes gene (w^+), which appeared to be repressed when it was translocated from its normal euchromatic environment to within the close proximity of heterochromatin. The result of this translocation was phenotypically manifested as a transition from white eyes to a variegated white and red eye color phenotype in cells that had repressed w^+ due to the heterochromatic position effect.

Heterochromatin is generally classified as either constitutive or facultative, both of which can repress gene activity. Constitutive heterochromatin is a highly ordered structure that is associated with distinct chromosomal domains, namely centromeres, their neighboring pericentromeric regions, and telomeres (Craig 2005). Centromeres form transcriptionally silent compartments in the nucleus, and can repress active genes that are relocated either within them or their vicinity (Grewal and Elgin 2002; Elgin and Grewal 2003). Nevertheless, constitutive heterochromatin domains are not static and do not always lead to permanent repression of genes (Dillon 2004). In contrast, facultative heterochromatin does not appear to have a highly regular structure, is more accessible than

constitutive heterochromatin and is developmentally regulated to silence the inactive X chromosome (X_i). Hallmarks of both constitutive and facultative heterochromatin are hypermethylation of DNA, hypoacetylation of histones, and hypermethylation of histone residues at K27-H3, K9-H3 and K20-H4, albeit of distinct modification states (mono-, di-, or tri-Me) for both heterochromatin subtypes. Importantly, histone methyl-lysine marks in heterochromatin are thought to be quite stable and therefore have the potential to mediate long-term repression throughout development.

Pericentromeric heterochromatin

Pericentromeric heterochromatin forms prominent domains in the nucleus that can be visualized using dyes such as DAPI or Hoechst that bind AT-rich DNA, which is enriched in satellite repeats (Nielsen et al. 2001; Schotta et al. 2004). The formation of these domains is one of the best characterized processes in histone methyl-lysine biology and appears to involve an enrichment in mMeK27-H3, tMeK9-H3 and tMeK20-H4 (Peters et al. 2003; Rice et al. 2003; Schotta et al. 2004). tMeK9-H3 is catalyzed by Suv39h1/2 HKMTases and generates a high affinity binding site for HP1, which further stabilizes and propagates heterochromatin formation (Schotta et al. 2003). Furthermore, the rate at which pericentromeric heterochromatin is generated appears to be dependent on the levels of Suv39h1/2 (Ebert et al. 2004). tMeK9-H3 is thought to precede tMeK20-H4, which is catalyzed by Suv4-20h HKMTases (Schotta et al. 2004). The mechanism of mMeK27-H3 deposition in pericentromeric heterochromatin is not known although d/tMeK27-H3 can be catalyzed by Ezh2 (Ebert et al. 2004). Although tMeK9-H3 is required for downstream formation of tMeK20-H4 in pericentromeric heterochromatin and these modifications have

a striking overlap, the functional relationship between them is not clear. Finer mapping of both of these modifications indicate they are enriched at distinct repetitive elements (Martens et al. 2005). This suggests that tMeK9-H3 and tMeK20-H4 could be spatially distinct, despite their apparent overlap seen by immunofluorescence in pericentromeric heterochromatin. Because *Suv4-20h1/2*-deficient mice have not yet been reported it is difficult to address the potential differences in physiological functions of tMeK9-H3 and tMeK20-H4. This knowledge would be particularly important in view of our findings described in chapter 3, showing distinct distributions of these modifications during embryogenesis.

Telomeres

Telomeres appear to be structurally similar to pericentromeric heterochromatin in that they are also enriched in tMeK9-H3 by the action of Suv39h HKMTases and are bound by chromobox proteins that are homologous to HP1 (Garcia-Cao et al. 2004). Significant insight into the role of histone methylation in telomeres has come from mouse loss of function studies. In *Suv39h* double null (dn) cells, telomeres are elongated and are depleted in d/tMeK9-H3 and chromobox protein binding, but have an increase in mMeK9-H3. Mouse triple knockout cells (TKO) of the Retinoblastoma 1 (RB1) family, deficient in RB1, RBL1 and RBL2, have a similar phenotype to *Suv39h* dn cells with respect to chromosome segregation defects and abnormal telomere length but have normal levels of tMeK9-H3 in pericentromeric and telomeric heterochromatin (Gonzalo and Blasco 2005; Gonzalo et al. 2005). This result is surprising given the previously reported association between Suv39h and RB1 proteins (Ait-Si-Ali et al. 2004). Nevertheless, TKO cells have

depleted levels of tMeK20-H4 in constitutive heterochromatin but their levels of mMeK20-H4 are comparable to wild type (Gonzalo et al. 2005). It is therefore noteworthy that RB1 proteins can also interact with Suv4-20h HKMTases to regulate the formation of constitutive heterochromatin (Gonzalo et al. 2005). These findings are important in establishing that RB1-dependent tMeK20-H4 deficiency can manifest a similar phenotype to *Suv39h* dn cells without affecting the integrity of tMeK9-H3. This would suggest that while tMeK9-H3 and tMeK20-H4 may have similar functions in constitutive heterochromatin formation, these modifications could be involved in distinct molecular pathways.

Facultative heterochromatin

Facultative heterochromatin found on the X_i is enriched in mMeK20-H4, dMeK9-H3 and tMeK27-H3, and the mechanistic basis underlying this methylation profile is now beginning to be understood (Plath et al. 2003; Silva et al. 2003; Kohlmaier et al. 2004; Okamoto et al. 2004). The Embryonic ectoderm development (Eed)-Ezh2 complex is responsible for catalyzing tMeK27-H3 at the X_i *in vivo* (Plath et al. 2003; Silva et al. 2003). dMeK9-H3 has also been reported to be catalyzed by Eed-Ezh2 *in vitro* (Czermin et al. 2002; Kuzmichev et al. 2002), but whether this occurs *in vivo* remains controversial (Cao and Zhang 2004). The depostition of mMeK20-H4 on the X_i remains unclear, but could be catalyzed by Pr-Set7, as this HKMTase has been shown to catalyze this modification (Xiao et al. 2005). It is interesting to note that Pr-Set7 has been shown to influence K20-H4 methylation over the course of the cell cycle because Ezh2 is known to regulate tMeK27-H3 of the X_i in a cell-cycle dependent fashion (Rice et al. 2002; Caretti et al. 2004;

Karachentsev et al. 2005). Although it remains to be shown whether Pr-Set7 has a direct effect on the X_i, the possibility it modifies the X_i in a cell-cycle dependent manner remains intringuing and is consistent with our observations of mMeK20-H4 dynamics discussed in chapters 3 and 4 of this thesis.

Initiation of heterochromatinization by RNAi

RNA interference (RNAi) is a mechanism of gene repression induced by double-stranded RNA, which can be heritable and may mediate its effects though heterochromatin formation pathways (Grishok 2005; Kavi et al. 2005b). A mechanistic understanding of RNAi-directed heterochromatin formation is largely derived from studies in fission yeast (Kavi et al. 2005a). In *S. pombe*, small RNAs that are associated with centromeric repeats have been found and are termed heterochromatic siRNAs (Reinhart and Bartel 2002). Double-stranded RNAs originating from these repeats are cleaved by the RNase III-type endonuclease Dicer (Dcr1). This in turn generates short interfering RNAs (siRNAs) ~ 21 nucleotides in length that are incorporated into a nucleoprotein complex termed RNA-induced initiator of transcriptional gene silencing (RITS), which targets cognate mRNAs for degradation (Bayne and Allshire 2005). The RITS complex is comprised of the chromodomain protein Chp1, Ago1 and Tas3 (Verdel et al. 2004). In a transcription-dependent manner, the RITS complex associates with centromeric repeats and initiates the methylation of histone H3-K9 and further spreading of heterochromatin through Swi6 (SWItching deficient transcription factor) (Volpe et al. 2002; Verdel et al. 2004; Schramke et al. 2005). Furthermore, fission yeast mutants with perturbed RNAi machinery have defects in chromosome segregation and heterochromatin formation (Bayne and Allshire

2005). Similar heterochromatin defects are also seen by disrupting RNAi pathway components in *Drosophila*, mice, and chicken cells (Pal-Bhadra et al. 2004; Bayne and Allshire 2005). This would suggest that there is some conservation in the RNAi-mediated heterochromatin pathway, perhaps indicating a RITS-like complex is present in mammals.

RNAi is a conserved process, and there is increasing evidence that it nucleates heterochromatin formation in mammals as shown in *S. pombe* (Bayne and Allshire 2005; Filipowicz et al. 2005). It appears that in mammals and plants, but not in yeast, RNAi directs *de novo* DNA methylation as a prerequisite to heterochromatinization (Wassenegger 2005). Dicer-deficient mouse ES cells have depleted levels of centromeric histone H3-K9 methylation, and have *de novo* DNA methylation defects. The mouse *de novo* DNA methyltransferase DNMT3b is similar to the plant *de novo* methylase DRM2, which is an essential player in RNAi mediated heterochromatin formation (Wassenegger 2005). Chick Dicer mutants have a deficiency in pericentromeric heterochromatin formation and chromosome segregation (Bayne and Allshire 2005). In human cells, it has been shown that MeK9-H3-mediated repression of the E-cadherin promoter is dependent on RNAi. Notably, histone methyl-lysines are deposited onto the inactive X following Xist RNA coating suggesting that RNA may be a common precursor to heterochromatinization in mammals. The above studies provide strong support for the role of RNAi in initiating heterochromatin formation in mammals although the details of the molecular pathways involved remain to be elucidated.

Interrelationships of DNA methylation and histone methylation

Crosstalk between DNA methylation and histone methylation is quite complex involving various aspects of gene regulation that are sometimes taxa-specific. With respect to heterochromatin formation, however, there are general features that should be noted. The most recognized of these involves transcriptional inhibition of CpG dinucleotides in gene promoters. This effect can be mediated through either the transcription machinery itself or through histone modifications (i.e. hypoacetylation and MeK9-H3), which ultimately influences chromatin structure (Lande-Diner and Cedar 2005). It appears that histone methylation and DNA methylation may reinforce each other to generate stable heterochromatin structures (Stancheva 2005). In mouse pericentromeric heterochromatin, catalysis of tMeK9-H3 by Suv39h HKMTases is required for DNA hypermethylation and may involve recruitment of DNA methyltransferases by HP1 (Lehnertz et al. 2003; Freitag and Selker 2005). Conversely, ES cells deficient for DNA methyltransferases (Dnmt3a/Dnm3b) do not have perturbed tMeK9-H3 in pericentromeric heterochromatin (Lehnertz et al. 2003). In addition, a Suv39h-independent pathway appears to be involved for recruiting DNA methyltransferases to repressive chromatin in minor satellite repeats (Lehnertz et al. 2003). Similarly, during X-innactivation, dMeK9-H3 is established prior to DNA methylation (Stancheva 2005). Thus, it appears that at least in mammals, MeK9-H3-mediated heterochromatin formation is required for DNA methylation. Despite this hierarchy, various lines of evidence suggest that both modifications can act in a feedback loop to reinforce each other (Freitag and Selker 2005).

The interrelationship between DNA and histone methylation may help to explain how repressive histone methylation marks are propagated though cell generations (Stancheva 2005). Following DNA replication, hemimethylated DNA can be reestablished by maintenance methyltransferases such as DNMT1, generating a platform for the recruitment of HKMTases that could modify newly deposited nucleosomes. A semi-conservative mechanism for the replication of nucleosomes has been proposed, whereby H3:H4 dimers would remain on the parental strand and would enable the recruitment of HKMTases to reestablish the histone methylation pattern (Tagami et al. 2004). However, this model is difficult to reconcile with earlier studies suggesting that semi-conservative nucleosome deposition does not occur (Weintraub et al. 1978; Jackson 1988). An alternative model has been suggested by Henikoff et al., whereby modified nucleosomes are randomly distributed to both strands following DNA replication. In this manner, an approximate modification pattern sets the stage for a more refined distribution that is established by replication-independent mechanisms (Henikoff et al. 2004). It is important to address the mechanisms by which histone modifications patterns are maintained following DNA replication, as this is essential for the inheritance of epigenetic information during development.

Interplay between histone modifications

On a given histone tail, there appears to be the potential for complex interplay between various histone modifications due to the coexistence of adjacent modifications or competition for individual amino acids (Fischle et al. 2003) (Figure 1-2). Given the number, type and combination of modifications possible, only salient features of this

interplay will be mentioned here for simplicity. The histone H3-K9 residue represents a classic example of a site that may be modified by nearby residues, which could have antagonistic action. When acetylated, histone H3-K9 is associated with transcriptionally competent chromatin states and prevents heterochromatic methylation from occurring at this site and, vice versa. Trimethylation of histone H3-K4 appears to promote acetylation of adjacent lysines on the H3 tail - such as acK9-H3 and acK14-H3 – and the converse is true. When considering the interplay of several covalent modifications occurring between several residues on a given histone tail, this subject is clearly complex.

Lessons from HKMTase-deficient mice

Mouse HKMTase knockouts suggest that histone methylation is crucial for developmental regulation occurring from peri-implantation to post-natal life. These HKMTase-deficient mice have revealed distinct developmental roles for HKMTases that have appeared to have similar molecular functions *in vitro*. Although these studies demonstrate the importance of histone lysine methylation during embryogenesis, the substrate specificities of numerous HKMTases remain unknown. In addition, the spatial and temporal distributions of histone methyl-lysine derivatives have not been examined in either normal or HKMTase-deficient mice. Nevertheless, mouse HMTase knockouts have been informative in demonstrating the importance of these genes in mammalian development. Findings from mouse HKMTase knockouts generated to date will be discussed in this section.

Enhancer of Zeste 2 (*Ezh2*) is broadly expressed during embryogeesis as early as E6.5, its expression decreases by E11.5 and at E15.5 appears to be restricted to the thymus (O'Carroll et al. 2001; Caretti et al. 2004). *Ezh2* null mice have impaired outgrowth at the

blastocyst stage and cease to develop following post-implantation resulting in embryonic lethality as early as E5.5 (O'Carroll et al. 2001). A subset of *Ezh2*-deficient mice die later in development with defects resembling those of Embryonic ectodermal development (Eed) mutants; these embryos do not complete gastrulation as a result of widespread failure in morphogenesis (Faust et al. 1998). The parallel phenotypes between Ezh2 and Eed nulls suggests a common function and, not surprisingly, Ezh2 and Eed form a complex that regulates tMeK27-H3 of imprinted genes and random X-inactivation following Xist coating (Plath et al. 2003; Silva et al. 2003). A cell cycle defect is also apparent in *Ezh2* nulls, characterized by very short cycles, which supports a more global role for Ezh2 in cell proliferation (O'Carroll et al. 2001). However, the early lethality of *Ezh2* nulls has impaired characterization of its role in later development, a limitation that has been overcome through the generation of conditional knockouts. An *Ezh2* conditional knockout was designed to understand its elevated expression in lymphopoeisis and suggested a role in B cell development possibly through immunoglobulin heavy chain rearrangement, an essential process for the generation of specific antibodies (Yamane et al. 2006). Interestingly, histone H3-K27 methylation by Ezh2 is required for this process to occur.

Recent studies examing the role of Ezh2-mediated histone H3-K27 in the context of Polycomb Repressive Complexes (PRC) has lead to exciting findings regarding their function in developmental regulation of embryonic stem (ES) cells (Bernstein et al. 2006; Boyer et al. 2006; Bracken et al. 2006; Lee et al. 2006). Genome-wide analysis of PRC binding sites suggests these complexes may repress numerous developmental regulatory genes in ES cells through tMeK27-H3 in order to maintain pluripotentcy. In this model,

PRCs would withdraw from their repressive binding actions on developmental regulator genes. These new findings will likely lead to interesting insights into the epigenetic mechanisms of ES cell development.

Similar to Ezh2, Nsd1 has a crucial role in early post-implantation development and it appears to regulate cell proliferation. *Nsd1*-deficient embryos have severe growth retardation defects and display elevated apoptosis as early as E6.5 (Rayasam et al. 2003). *Nsd1* nulls fail to undergo normal gastrulation and by E10.5 *Nsd1* -/- embryos are resorbed. The broad expression profile of *Nsd1* suggests it has a more ubiquitous role in early development, although its expression is restricted to highly proliferative cells types by E14.5 (Rayasam et al. 2003). Importantly, this later tissue-restricted expression is consistent with the overgrowth syndrome observed in humans with *NSD1* deficiency (detailed in next section). Given the specificity of NSD1 for histone H3-K36 and H4-K20, it would be interesting to determine whether the presence of both modifications on a nucleosome has a discrete function in development, or whether the modifications are deposited independently by NSD1 in order to mediate distinct functions.

Suv39h1 and Suv39h2 appear to be globally co-expressed during mouse development, shown by their diffuse *in situ* localization in E8.5-9.5 mouse embryos (O'Carroll et al. 2000). Both HKMTases seem to have indistinguishable roles in catalyzing the formation of pericentric heterochromatin, while Suv39h2 may have an additional function in male meiosis (O'Carroll et al. 2000). Loss of either *Suv39h1* or *Suv39h2* does not influence viability or fertility, but mice lacking both genes display growth defects ranging in severity from lethality at E12.5 to reduced size in postnatal development, infertility and lymphoma

(Peters et al. 2001). Cells derived from *Suv39h* dn embryos were shown to have genomic instability, which suggested that pericentromeric heterochromatin is necessary for normal chromosome segregation. Importantly, the phenotype of *Suv39h* dn mice could also suggest that the integrity of pericentromeric heterochromatin does not have a significant effect on embryogenesis until mid-gestation development. Alternatively, other HKMTases may be compensating for Suv39h deficiency in early development, but other known histone H3-K9 HKMTases are not tMeK9-H3-specific and do not function in heterochromatin establishment.

Studies comparing the targeted loss of *G9a* and *Suv39h* dns were essential to demonstrate that although both HKMTases catalyze histone H3-K9 methylation, they have distinct functions due to their ability to generate different K9-H3 methyl states. This was first suggested by their disparate subnuclear targeting: Suv39h1/2 is enriched in pericentromeric heterochromatin (Bannister et al. 2001; Lachner et al. 2001), while G9a associates with euchromatic regions (Tachibana et al. 2001). In contrast to *Suv39h* dns, *G9a* deficient mice have severe growth retardation defects that results in embryonic lethality during gastrulation (Tachibana et al. 2002). It appears that while G9a is not necessary for the survival and proliferation of undifferentiated cells, it is important for somatic differentiation processes. Consistent with a role for G9a in euchromatic histone H3-K9 dependent repression, *G9a -/-* cells have a loss of dMeK9-H3 and a corresponding increase in MeK4-H3 and acK9-H3 (Tachibana et al. 2002). Further analysis of *Suv39h* dn and *G9a* deficient cells solidified their distinct functions by showing that G9a catalyzes mono- and dimethylation whereas Suv39h1/2 are required for trimethylation (Rice et al.

2003). It would be valuable to generate *Suv39h* dn/*G9a* triple knockout embryos to determine which cellular processes may be independent of these HKMTases in early development.

ESET is similar to G9a in that it is a euchromatic K9-H3 methylase and is also required for early mouse development (Dodge et al. 2004). *Eset* appears to be ubiquitously expressed throughout mouse embryogenesis and, null mice die during peri-implantation between E3.5 and E5.5, just slightly earlier than *G9a* mutants. In addition, *Eset* mutants do not appear to have altered tMeK9-H3 levels supporting the role of ESET as a euchromatic K9-H3 methylase. Further analysis of *Eset* in comparison to *G9a* mutants using conditional knockout strategies would be interesting in order to determine their functional differences *in vivo*.

Several Mll loss-of-function mouse models have been examined to understand its role. *Mll* is broadly expressed during embryogenesis from E7.0 onward and its targeted disruption results in gross morphogenetic defects that begin to appear in mid-gestation development (Yu et al. 1995). By E9, Mll -/- mice have a loss of *Hox* gene expression and by E10, gross malformations of cranial and spinal ganglia, and loss of somite boundaries are observed. Importantly, Yu et al. suggested that E8.5-E9 is a critical period in embryogenesis dependent on *Mll* and hypothesized that this interval in development may be "a dynamic period of chromatin reorganization, which leads to either silencing or persistent expression of developmentally regulated loci". With respect to histone lysine methylation, this developmental window could be marked by genome-wide changes in histone modifications, a concept that is pursued in this thesis.

Given the widespread involvement of *Mll* in leukemia (detailed in next section), its effect on hematopoiesis was examined by generating *Mll*-deficient ES cells, which can be cultured under certain conditions to induce differentiation towards hematopoietic lineages (Fidanza et al. 1996). This study indicated that Mll is required for hematopoietic maturation and, using a similar loss-of-function assay, Ersnt et al. showed that Mll-targeted *Hox* genes are essential for hematopoiesis, solidifying the role of Mll in *Hox* gene regulation (Ernst et al. 2004b). Using *Mll* mutant chimeras, it was shown that *Mll* deficient cells have a severe block in definitive hematopoiesis, and demonstrated an essential role for Mll in hematopoietic stem cell generation (Ernst et al. 2004a). These findings provide an important example of the role of histone methylation in differentiation.

Ectopic viral integration site 1 (Evi1 or Prdm3) is expressed at high levels as early as E8.5 in primary headfolds and by E9.5-10.5 extends to several other tissues including the forebrain, nasal placode, first and second branchial arches, mesonephros, limb buds and dorsal root ganglia (Hoyt et al. 1997). *Evi1*deficient mice die at E10.5 from a complex phenotype affecting multiple organs with the main cause of lethality attributed to compromised circulation from an underdeveloped heart. By E10.5, gross underdevelopment of the nervous system is apparent with a lack of spinal neurons. In the somites of E10.5 knockout embryos, the dermatome appeared normal whereas the myotome and sclerotome were grossly acellular and, consistent with this defect, limb buds were underdeveloped. Moreover, subtle malformations were seen in the developing liver, mesonephros and branchial arches. The widespread hypocellularity in *Evi1* mutants would suggest a general role for Evi1 in proliferation, perhaps related to its involvement in

leukemia (Nucifora 1997). In addition, the specific organ defects seen in these mutants and the elevated *Evi1* expression in distinct tissues would indicate a greater requirement for Evi1 in certain developmental lineages. In this regard, it would be valuable to determine histone modification specificity of Evi1, and whether this modification plays a cell lineage-specific function.

Deregulation of histone methyltransferases in human disease

Carcinogenesis

In normal cells, a precise balance between euchromatin and heterochromatin needs to be stably maintained, which is achieved by the combinatorial action of histone modifications and DNA methylation (Jones and Baylin 2002). Several examples have demonstrated that deregulation of this balance results in malignancy. DNA methylation was historically the first epigenetic modification found to have an involvement in cell transformation and it is now known to have a pivotal role in the predisposition and progression of human cancers (Frigola et al. 2005; Scarano et al. 2005). Both hypomethylation and hypermethylation of CpGs can be tumorigenic, by aberrantly activating or repressing key regulatory genes, respectively. For instance, promoter hypermethylation of the DNA mismatch repair gene *MLH1* can result in colorectal, endometrial and gastric cancers (Esteller 2005), whereas hypomethylation of *HRAS* and *cMYC* enables expression of these transforming oncogenes (Scarano et al. 2005). Hypomethylation of DNA can also result in genome instability by predisposing repetitive regions of the genome to chromosomal rearrangements. Several examples of deregulated DNA methylation events described above have been shown and in

many instances this epigenetic mark has diagnostic and prognostic value (Scarano et al. 2005). Despite the incredible progress that has been made in characterizing the role of DNA methylation in carcinogenesis, it is important to note its interdependence with other epigenetic marks, such as histone post-translational modifications. Recently, histone lysine methylation has been shown to have a particularly important role in carcinogenesis, which is demonstrated by the fact that to date approximately 1/3 of HKMTases are implicated in human malignancy (Schneider et al. 2002). This is not surprising, as many HKMTases have been shown to play an important role in cell growth control. Although the mechanisms by which HKMTases give rise to cancer remain elusive in several instances, the carcinogenic role of some HKMTases have been studied in detail and will be reviewed here.

PR domain containing proteins are tumor suppressors

To date, all PR domain containing genes that have been characterized have tumor suppressive properties and several members are disrupted in cancer (Deng and Huang 2004). PRDM genes also have remarkably similar regulatory mechanisms that are disrupted in cancerous cells. For instance, RIZ1 (Prdm2), Prdm1, and Evi1 (Prdm3) have alternative promoter usage that generates 2 functional mRNAs that differ only by the presence of the PR domain (Derunes et al. 2005). The PR domain containing products are anti-tumorigenic, whereas the PR-deficient forms are generally overexpressed in tumors and behave as oncogenes (Nitta et al. 2005).

RIZ1 was the first HKMTase shown to function as a tumor suppressor (discussed further in next section) and is the best characterized PR domain gene (Huang et al. 1998;

Huang 2002). *RIZ1* is located at 1p36, a site that is commonly translocated in cancer, resulting in production of the PR-domain protein *RIZ2* (Huang 2002). Aberrant silencing of *RIZ1* can give rise to neuroblastoma, melanoma, osteosarcoma and cancers of the lung, breast, colon, and liver (Huang 2002). Missense mutations near or within the *RIZ1* PR domain also result in cancer, some of which disrupt residues that correspond to those required in Suv39h1/2 for its catalytic activity (Schneider et al. 2002). Interestingly, RIZ1 appears to share similar functional features with Suv39h1/2. *RIZ1* knockout mice develop tumors that are reminiscent of *Suv39h* dns and, like Suv39h proteins, RIZ1 interacts with Rb and is a histone H3-K9 HKMTase (Huang et al. 1998; Huang 2002; Kim et al. 2003). Importantly, administration of RIZ1 to colon cancer xenografts has anti-tumorigenic effects that may hold promise for future gene therapy (Canote et al. 2002).

Evi1 has two protein products, the PR-deficient form (Evi1a) and the full-length product Evi1c (MDS1-Evi1). Overexpression of Evi1a can occur by viral translocation or chromosomal rearrangement and results in acute myeloid leukemia with unfavorable prognosis (Valk et al. 2004). Evi1a can also repress the growth inhibitory effects of transforming growth factor-β (TGF-β) and inhibit differentiation of various cell types (Nitta et al. 2005). In addition, Evi1a can interact with a corepressor protein CtBP to further repress TGF- β signaling. In contrast, Evi1c cannot repress the TGF-β pathway unless it is is fused to the Acute Myeloid Leukemia 1 gene (AML1-Evi1) where it is pathogenic and confers similar properties to Evi1a (Izutsu et al. 2002). Both AML1-Evi1 and Evi1a, but not Evi1c, can homo-oligomerize, a property that is thought to be important for Evi1-mediated leukemogenesis (Nitta et al. 2005).

MLL

Mixed lineage leukemia (MLL) has been extensively studied since it was first isolated as a prevalent target of chromosomal translocation in human leukemia. To date, MLL has been associated with greater than 60 chromosomal rearrangements and 33 MLL fusion proteins have been characterized (Daser and Rabbitts 2004; Forneris et al. 2005). Given the requirement of MLL for definitive lymphopoeisis, it is not surprising that its deregulation can give rise to leukemia. MLL rearrangements manifest various leukemogenic subtypes through different mechanisms (Forneris et al. 2005). Perhaps the most common oncogenic pathway involves fusions to transcriptional activator proteins. The result is aberrant transactivation that varies in degree depending on the heterologous activation domain attached to MLL. MLL fusions also occur with cytoplasmic proteins that have self-association motifs, thereby generating oligomerized chimeras that behave as aberrant transcription effectors. In addition, MLL fusions can occur with SH3 domain proteins resulting in deregulated signaling pathways that are leukemogenic. Deregulated gene expression caused by MLL fusions is likely to be largely attributed to MLL-mediated chromatin changes. Because MLL is a histone H3-K4 HKMTase, its mistargeting could lead to aberrant gene expression, which could happen through protein interactions mediated by the MLL fusion partner. Consistent with such a scenario, MLL fusion proteins have been shown to interact with SWI/SNF chromatin remodeling proteins (Hess 2004).

Deregulation of NSD gene family members

Nuclear SET Domain (NSD) HKMTases NSD1, NSD2 and NSD3 form a small gene subfamily with highly similar gene structure comprising a PHD (plant homedomain),

PWWP (proline-tryptophan-tryptophan-proline), and SET domain, and all three members have been implicated in human carcinogenesis. In some cases of childhood acute myeoloid leukemia (AML), *NSD1* is involved in a translocation that generates a fusion protein with *Nup98*, a component of the nuclear pore complex (Cerveira et al. 2003). In other hematological malignancies, *Nup98* is known to be fused to at least 11 other genes, but oncogenic mechanisms involving *Nup98* fusions are not clear (Panagopoulos et al. 2003). Happloinsufficiency of *NSD1* is associated with Sotos syndrome (SS) (detailed in next section), a disorder that appears to have a predisposition to malignancy. To date, 22 SS patients (~2-3%) have developed malignant tumors, which typically manifest during childhood (Al-Mulla et al. 2004; Deardorff et al. 2004; Lapunzina 2005). *NSD2* (also known as *MMSET* or *WHSC1*) is implicated in myelomagenesis where it is overexpressed in $t(4;14)^{pos}$ translocation patients (Keats et al. 2005). *NSD2* is transcribed to generate three alternatively spliced species (MMSET I-III) in addition to a second product that originates within *NSD2*, termed REIIBP. Interestingly, breakpoint variants localize to disparate regions of the nucleus and exhibit different kinetic properties when measure by fluorescence recovery after photobleaching (FRAP), possibly reflecting their interaction with distinct chromatin compartments. Future studies aimed at dissecting the complexities of *NSD2* function will likely contribute to a better understanding of myelomagenesis. It appears that *NSD3* may also be deregulated in cancer as it is located within a region that is amplified in breast carcinoma (Angrand et al. 2001). Similar to *NSD1*, a *NSD3* translocation event that generates a fusion protein with Nup98 has been identified in a patient with AML (Rosati et al. 2002).

H4K16 acetylation and K20-H4 methylation

A recent study examining the influence of histone H4 modification in human cancer deserves special attention, as it is particularly relevant to this thesis. Fraga et al. demonstrated that acK16-H4 and tMeK20-H4 are significantly depleted in repetitive elements in leukemia cells and several other tumor types, including breast, colon and lung cancer (Fraga et al. 2005). The authors propose that the loss of these modifications may be a common feature of malignant transformation, akin to global DNA hypomethylation and CpG island hypermethylation. Given the association tMeK20-H4 with gene repression, it is speculated that its loss in cancer cells results in a gain-of-function of cell-cycle associated genes and, as such this modification would be required for cellular differentiation events. As will be demonstrated in the following chapters, tMeK20-H4 is in fact associated with specific differentiation events in various cell types.

Developmental disorders

NSD1 in childhood overgrowth syndromes

Sotos syndrome (SS) is one of several childhood overgrowth syndromes, which is clinically delineated by distinctive facial features, accelerated growth, developmental delay and a variable spectrum of associated abnormalities (Baujat et al. 2005). The majority of SS cases are caused by mutations in *NSD1*, which have been identified in more than 150 patients to date (Faravelli 2005). In SS patients, *NSD1* is most often disrupted by intronic mutations that disrupt transcription or missense mutations in conserved domains, in addition to less frequent microdeletions at 5q35 (Douglas et al. 2005; Tatton-Brown et al.

2005a; Tatton-Brown et al. 2005b). A variable number of SS patients do not have *NSD1* mutations and this may be due to allelic or locus heterogeneity (Faravelli 2005). In an attempt to identify other SS loci, Douglas et al. screened several patients with overgrowth phenotypes similar to SS for mutations in *NSD2* and *NSD3*, both of which share considerable similarity with *NSD1* (Douglas et al. 2005). This screen did not reveal any pathogenic mutations in either gene, suggesting they operate in molecular pathways that are distinct from *NSD1*. Although downstream targets of *NSD1* have not yet been identified, they may be responsible for SS cases not involving *NSD1* mutations.

NSD1 mutations have been less frequently identified in other overgrowth syndromes that have phenotypic overlap with SS, namely Weaver syndrome (WS) and Beckwith-Wiedemann syndrome (BWS) (Baujat et al. 2005). Clinically distinctive features segregate WS and SS, but the presence of *NSD1* mutations in both syndromes indicates the conditions are allelic. BWS is caused by deregulation of imprinting at 11p15, resulting in overexpression of IGF2. Interestingly, *NSD1* mutations have been identified in 2 patients with BWS and 11p15 uniparental disomy has been found in 2 cases of atypical SS (Baujat et al. 2004). In this regard, one could determine whether *NSD1* may regulate genes at 11p15 or vice versa. At the molecular level, there is currently limited knowledge about the function of NSD1, making it difficult to assess how it could contribute to the pathophysiology of overgrowth syndromes. NSD1 appears to have dual specificity for MeK36-H3 and MeK20-H4, which are generally marks of active and repressive chromatin, respectively. It would be interesting to determine how alterations in these modifications could lead to overgrowth syndromes.

Wolf-Hirschhorn Syndrome

Wolf-Hirschhorn syndrome (WHS) is characterized by distinct craniofacial anomalies, epilepsy and growth delay in addition to a variable spectrum of secondary defects. The *NSD2* gene is invariably deleted in this syndrome due to microdeletions at 4p that vary in size (Bergemann et al. 2005). It is thought that the pleiotropic secondary phenotypes seen in WHS could be a result of deregulated expression of NSD2 target genes. In addition, other candidate genes that span WHS 4p deletions may have a role in manifesting secondary defects. Further studies are required in order to get a molecular understanding of WHS.

9q34 subtelomeric deletion syndrome

Euchromatin Histone Methyltransferase1 (Eu-HMTase1) is a K9-H3 HKMTase similar to G9 (Ogawa et al. 2002) and has recently been implicated in 9q34 subtelomeric deletion syndrome (9q34 SDS) (Kleefstra et al. 2005). Patients with this newly designated syndrome have subtelomeric deletions at 9q and present with mental retardation, a range of craniofacial malformations and heart defects (Kleefstra et al. 2005). Eu-HMTase1 is disrupted by translocation in a patient with a clinical profile matching 9q34 SDS and therefore appears to be causative in this syndrome. In addition, the expression of Eu-HMTase *in silico* and in developing mice is consistent with 9q34 SDS. Eu-HMTase1 knockout mice have not yet been generated, but would be informative in understanding its developmental function and pathogenesis in 9q34 SDS.

Cardiac disease

Elevated plasma homocysteine is a significant risk factor in cardiovascular disease, however the basis of this is not clear (Moat et al. 2004a; Moat et al. 2004b). Dietary supplementation with folic acid in conjunction with vitamins B6 and B12s has been shown to ameliorate endothelial dysfunction and hypercholesterolemia (Verhaar et al. 1999; van Etten et al. 2002). Interestingly, elevated homocysteine is associated with a decrease in SAM, the methyl donor for HKMTase reactions (Huang 2002). It may therefore be possible that homocysteinuria impedes HKMTase reactions. Recently, the HKMTase Smyd1 (m-Bop) has been shown to have a specific role in cardiac development (Gottlieb et al. 2002; Phan et al. 2005) and is implicated in end stage heart disease (Borlak and Thum 2003). It is also interesting to note that *Smyd1* is regulated by MEF2C, which in turn is regulated by $Ca2^{+}$-dependent signal transduction involving calreticulin and calcineurin (Lynch et al. 2005; Phan et al. 2005). Because both calreticulin and calcineurin have been implicated in cardiac hypertrophy (Wolk 2003; Lynch et al. 2005), it is possible that Smyd1 may have a role in this cardiac pathology as an epigenetic effector in the calreticulin/calcineurin pathway. In addition to histone methylation, histone acetylation has been shown to play a role in cardiac disease, perhaps as a result of the interdependence of these modifications. HDACs 4, 5, 7 and 9 expression is elevated in heart and these enzymes cooperate with cardiac-specific transcription factors such as MEF2 and Calmodulin. HDACs 5 and 9 have been implicated in cardiac hypertrophy, a condition that can be repressed by HDAC inhibitors (McKinsey and Olson 2004). Further dissection of the roles of histone modifications in the heart will likely contribute to a better

understanding of such insidious cardiac diseases and could provide additional therapeutic targets.

Objectives and Rationale

An understanding of the role of histone lysine methylation in development is in its infancy. The *in situ* distributions of several histone methyl-lysines are unknown and their dynamics in developmental processes remain to be examined throughout embryogenesis. In addition, the developmental and cancer predisposing roles of HKMTases remain largely elusive. These areas of histone methyl-lysine biology leave interesting and important questions to be explored.

To begin to address the distributions of histone methylation during development, we examined several methyl-lysine derivatives at the cellular and tissue level during mid-gestation in mammalian embryos. At the time we began this project, these distributions had not been described, leaving a gap in knowledge between involvement of HKMTases in development and potential changes in histone lysine methylation profiles occurring during these processes. We reasoned that if broad changes in chromatin domains occur during development, a mechanism for the stable epigenetic propagation of these changes should involve histone lysine methylation. Indeed, we found that m/tMeK20-H4 underwent marked changes over the course of neuronal and myogenic development (Chapter 3).

We aimed to further characterize histone methyl-lysine profiles during skeletal muscle differentiation, which offers a well-charaterized and tractable model system. This investigation uncovered other histone methyl-lysine residues that have dynamic properties during differentiation, namely mMeK36-H3, m/tMeK79-H3 and m/d/tMeK20-H4 (Chapter

4). We then sought to identify the HKMTases responsible for the histone methyl-lysine dynamics observed in myogenesis (Chapter 5). To this end, we have begun to characterize the functions of Prdm12 and Smyd1 in muscle differentiation. Together, our analysis has provided fundamental knowledge about the role of histone lysine methylation during development.

Figure 1- 1: Subnuclear compartments

The mammalian nucleus is organized into distinct domains. Nucleoli are the largest nuclear domains and are involved in ribosome biogenesis. Sam68 bodies and the perinucleolar compartment are found at the periphery of nucleoli. Heterochromatin is highly ordered and is generally distributed at the nuclear and nucleolar periphery, whereas euchromatin is less compact and found within the nucleoplasmic interior. Highly active chromatin is enriched at the periphery of splicing speckle compartments, which contain pre-mRNA splicing factors. Paraspeckle proteins are also associated with the splicing speckle periphery. PML, OPT and PcG domains vary in number and location throughout the nucleoplasm. Cleavage bodies, Cajal bodies and Gems are often found in association with one another, and are also commonly associated with the nucleolar periphery.

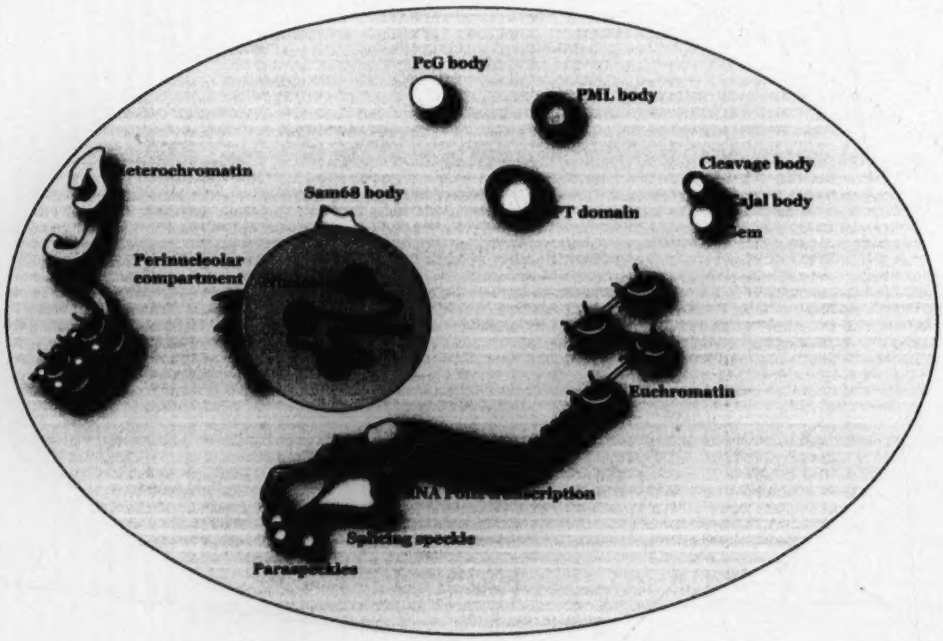

PcG body
PML body
Cleavage body
Sam68 body
T domain
Perinucleolar compartment
Euchromatin
Splicing speckle
Paraspeckles

Figure 1- 2: Histone acetylation, phosphorylation and methylation.

Histone tails, shown protruding from the nucleosome, are subject to numerous post-translational modifications including acetylation (Ac), phosphorylation (P) and methylation (Me). Lysines can be mono-, di- or tri-methylated, whereas arginines can be mono- or di-methylated in symmetric or assymetric conformations. Additional histone modifications not shown here include ADP ribosylation, ubiquitination and citrullination, which are also thought to contribute to chromatin function, but are not as well characterized.

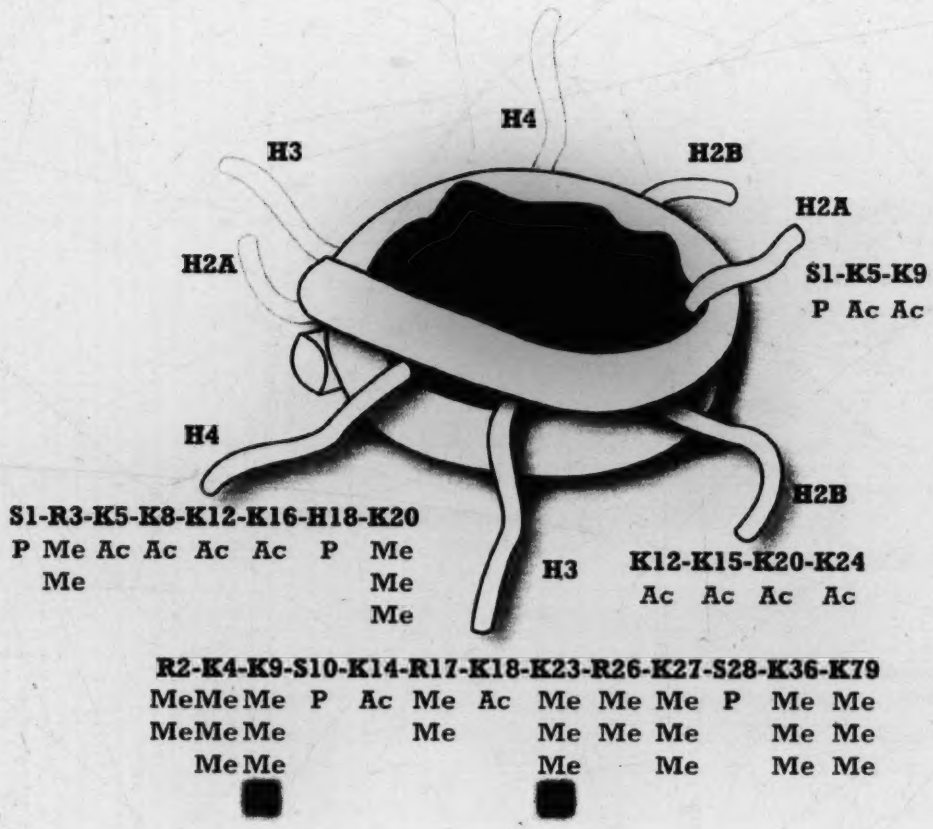
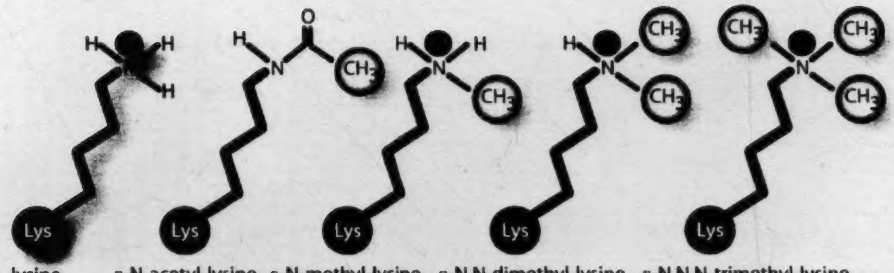
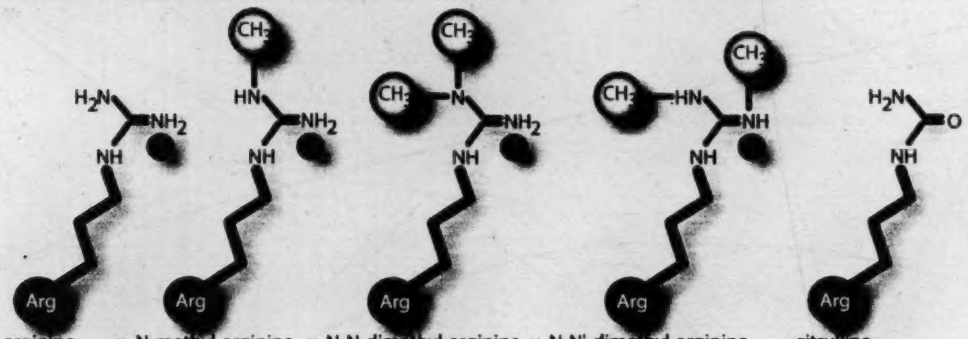

H4
H3
H2B
H2A
H2A
S1-K5-K9
P Ac Ac
H4
H2B
S1-R3-K5-K8-K12-K16-H18-K20
P Me Ac Ac Ac Ac P Me
Me
Me
Me
H3
K12-K15-K20-K24
Ac Ac Ac Ac
R2-K4-K9-S10-K14-R17-K18-K23-R26-K27-S28-K36-K79
MeMeMe P Ac Me Ac Me Me Me P Me Me
MeMeMe Me Me Me Me Me Me Me Me
MeMe Me Me Me

Figure 1- 3: Histone modification chemistry.

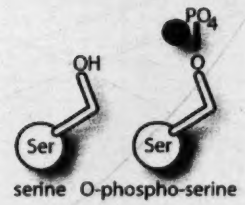
The amino group of histone lysines can be acetylated, which neutralizes the positive charge and results in repulsive electrostatic interactions with negatively charged DNA. Histone lysines can also be or mono-, di- or tri-methylated, conserving most electrostatic interactions, but adding Van der Waals forces. Arginines can be mono- or di-methylated, or citrullinated and di-methyl arginines may occur in symetric or assymetric conformations. By removal of the amino group, methyl arginines can be citrullinated. Serine residues can be phosphorylated as shown.



lysine ε-N-acetyl-lysine ε-N-methyl-lysine ε-N,N-dimethyl-lysine ε-N,N,N-trimethyl-lysine



arginine ω-N-methyl-arginine ω-N-N-dimethyl-arginine asymetric ω-N-N'-dimethyl-arginine symetric citrulline



serine O-phospho-serine

Experiments involving animals were approved by the Health Science Animal Policy Welfare and Committee (HSAPWC) at the University of Alberta.

Immunofluorescence

Mouse embryo immunofluorescence

Pregnant CD1 mice obtained from Charles River Laboratories were sacrificed by cervical dislocation and E8.5, E9.5, E10.5, and E11.5 embryos were collected by cesarean. Embryos were washed in PBS and fixed in 4% paraformaldehyde (PF) at 4°C for 18 hrs, cryoprotected in 30% sucrose-PBS, mounted in O.C.T. (Tissue-Tek) and sectioned at 14 µm using a Leica CM 1900 cryostat. Sections were immediately fixed in 4% PF for 7 min, washed 2 x 10 min in PBS, permeabilized for 30 min in PBS/0.1% Triton X-100 (P/T), blocked in P/T/5% heat inactivated sheep serum (P/T/S) for 30 min, and incubated overnight at 4°C with primary antibody (rabbit-α-mono-methylK4-H3, 1/500; rabbit-α-tri-methylK4-H3, 1/800; rabbit-α-tri-methylK20-H4, 1/1500; rabbit-α-mono-methylK20-H4, 1/1500; rabbit-α-tri-methylK9-H3, 1/750; rabbit-α-mono-methylK9-H3, rabbit-α-phosphoS10+tri-methylK9-H3 1/1000 [AbCam]; mouse-α-MF20, 1/10 supernatant [Developmental Studies Hybridoma Bank]; mouse-α-phosphoS10, 1/50 [Upstate Biotech]). Methylation antibodies were competed effectively by the immunogen but not by peptides containing different methyl modifications of the same lysine or an equivalent degree of methylation at a distinct position [AbCam]. Embryo sections were washed 3 x 10 min in P/T, blocked for 30 min in P/T/S, incubated with secondary antibody (Alexa594-

goat-α-rabbit, 1/300 [Molecular Probes]) for 2 hrs at room temperature, washed 3 x 5min in P/T and mounted with 90% glycerol-PBS-DAPI (1μg/ml).

Chick embryo immunofluorescence

O.C.T. (Tissue Tek) embedded sections from D5, D7 and D16 chick embryos were obtained from Dr. Roseline Godbout and Sachin Katyal (University of Alberta). Sections were fixed in 4 % paraformaldehyde for 2 minutes, washed in PBS (3 x 5 minutes), permeabilized in 1% IGEPAL CA630 (Sigma) in PBS for 5 minutes and washed in PBS (5 x 2 minutes). Slides prepared for PCNA staining were treated with 4M HCl for 10 minutes and washed in PBS (3 x 2 min) prior to blocking. Sections were blocked for 1 hour with 10 % milk and incubated overnight at 4°C with primary antibody diluted in 10% milk (rabbit-α-tri-methylK20-H4, 1/500; rabbit-α-mono-methylK20-H4, 1/500 [Abcam]; mouse-α-PCNA 1/100 [Sigma].) Slides were then washed in PBS (3 x 5 min) and incubated with secondary antibody for 2 hours at room temperature (mouse-α-Alexa 488, 1/200; rabbit-α-Alexa 594, 1/200), washed in PBS (3 x 5 minutes) and mounted with 90% glycerol:PBS containing Hoechst 33325 (1/1500).

Immunofluorescence with immortalized cultures

Mouse embryo fibroblast cells (C3H 10T1/2) were cultured in α-MEM + 10% FBS in a 37°C incubator with 5% CO_2 and were plated onto sterilized glass coverslips so that they were 50 to 80%-confluent the following day. The cells were permeabilized with PBS containing 0.5% Triton X-100 for 10 min and washed twice with PBS prior to serial incubations with 30 μl aliquots of an appropriate primary antibody (30 min at room

temperature) followed by a 30 min incubation with a Cyanin-3 conjugated anti-mouse secondary antibody (1:200 [Jackson ImmunoResearch Laboratories, Inc.]). Cells were rinsed with PBS containing 0.1% Triton X-100 and washed twice with PBS. Coverslips were mounted onto slides containing approximately 10 µl of a 90% glycerol-PBS-based medium containing 1mg/ml paraphenylenediamine and DAPI (0.5 µg/ml).

C2C12 cells were cultured in DMEM supplemented with 4 mM L-glutamine and 20% FBS in a 37°C incubator with 5% CO_2 and were plated onto sterilized glass coverslips so that they were 50 to 80%-confluent the following day. The cells were permeabilized with PBS containing 0.5% Triton X-100 for 10 min and washed twice with PBS prior to incubations with a 50 µl aliquot of an appropriate primary antibody (mouse-α-Prdm12 [Abcam], mouse-α-Pax3 supernatant, 1/10; mouse-α-Pax7 supernatant, 1/10; mouse-α-myogenin concentrate, 1/100 [DSHB]; mouse-α-Myod, 1/100 [DAKO]; mouse-α-BrdU, 1/50; rabbit-α-Flag M3, 1/100 [Sigma]; mouse-α-Myf5, 1/100 [Santa Cruz]; rabbit-α-mono-methylK4-H3, 1/500 [Abcam]; rabbit-α-di-methylK4-H3, 1/200 [Abcam]; rabbit-α-tri-methylK4-H3, 1/800 [Abcam]; rabbit-α-tri-methylK20-H4, 1/800 [Abcam]; rabbit-α-di-methylK20-H4, 1/800 [Abcam]; rabbit-α-di-methylK20-H4, 1/800 [Abcam]; rabbit-α-tri-methylK79-H3, 1/600 [Abcam]; rabbit-α-di-methylK79-H3, 1/600 [Abcam]; rabbit-α-mono-methylK79-H3, 1/600 [Abcam]; rabbit-α-tri-methylK9-H3 [Abcam], 1/750; rabbit-α-mono-methylK9-H3, 1/500 [Abcam]; rabbit-α-phosphoS10+tri-methylK9-H3, 1/1000 [AbCam]; rabbit-α-monomethylK36-H3, 1/500 [AbCam]; rabbit-α-trimethylK36-H3, 1/500 [AbCam]) for 45 min at room temperature followed by a 45 min incubation with a secondary antibody (α-mouse Alexa 594, 1/500; α-mouse Alexa488, 1/400; α-rabbit

Alexa594, 1/400 [Molecular Probes]). Cells were rinsed with PBS containing 0.1% TritonX-100 and washed twice with PBS. Coverslips were mounted onto slides containing approximately 10 µl of a 90% glycerol-PBS-based medium containing DAPI (0.5 µg/ml).

Immunofluorescence with primary cultures

Embryonic limb buds were dissected from E10.5 CD1 mice and prepared for tissue culture as previously described (Cash et al. 1997; Weston et al. 2000). Dissection instruments were washed with 70% ethanol and dried with Kimwipes prior to cesarian of sacrificed CD1 mice. Dissected embryos were placed in Puck's saline A (PSA) and whole limbs were removed with tweezers, minced into 2-4 tissue blocks/limb and transferred to 15 ml conical tubes, which were centrifuged at 1000 x g for 5 min. PSA was removed and replaced with 2.7 ml of PSA (Puck's saline A) containing 10% fetal bovine serum along with 0.3 ml of Dispase [Gibco](10 Units/ml). Limb bud tissue was incubated for ~1 hr at 37°C with slight shaking (~70 rpm) and gently vortexed every ~ 15 mins. Following incubation, 3 ml of culture media was added (40% DMEM/60% F12/ 10% FBS supplemented with penicillin, streptomycin and glutamine), and triturated. Cells were centrifuged at 1000 x g for 5 min, followed by removal of supernatant and the resuspension of cells in 5 ml of fresh media, which was strained though a pre-wetted 40µm cell strainer (Falcon 2340, cell strainer) to remove ectoderm and cell debris. An additional 3 ml of media was passed through the filter and the filtered cells were transferred to a 15 ml conical tube. Cell concentration was calculated using a Coulter counter and cells were then centrifuged at 1000 x g for 5 min. Supernatant was removed and cells were resuspended to a concentration of 1-2 X 10^7 cells/ml. Cells were then dispensed in 20 µl aliquots onto glass

coverslips in 6-well plates and incubated in a tissue culture incubator (37°C, 5% CO_2) for 1-2 hrs. Following this, 1 ml of media was added and replenished daily. Cells were fixed with 4% PF after 1, 5 or 7 days of culture and processed for immunofluorescence with the antibodies described above.

Bromodeoxyuridine labeling of primary cultures

Primary cultures were incubated for 2 hrs in primary culture media, as previously described (Cash et al. 1997; Weston et al. 2000), containing 5 mM bromodeoxyuridine (Sigma). Cells were fixed in 4 % paraformaldehyde for 10 min, washed in distilled water (2 x 1 min), permeabilized with acetone (-20 °C) for 2 min at room temperature and rinsed in water (5 x 30 s). DNA was denatured by adding 2N HCl for 15 min at room temperature and neutralized by washing 2 x 2 min with 0.1M sodium borate (pH 8.5). Cells were then washed in PBS for 5 min, blocked in 5 % BSA for 30 min, incubated with primary antibodies (mouse-α-BrdU, 1/50 [Sigma]; H4-monoMeK20, 1/500[Abcam]; H4-triMeK20, 1/500 [Abcam]), washed in PBS (3 x 5 min) and incubated with secondary antibodies (mouse-α-Alexa 488, 1/200; rabbit-α-Alexa 594, 1/200). Coverslips were mounted on slides as described above.

Analysis using Knockout mice

Calreticulin (*Crt*) null E10.5 embryos were obtained from Dr. Marek Michalak (University of Alberta) and were identified by PCR genotyping analysis of yolk sac DNA.

Yolk sac DNA extraction and genotyping

Yolk sac tissue was incubated in 40 µl of 25 mM sodium hydroxide/0.2mM EDTA (pH 12) at 95 °C for 20 minutes, 40 µl of 40 mM Tris HCl (pH 5) was added and samples were briefly vortexed and spun for 6 minutes at 12, 000 x g. Supernatant was transferred to a new tube and 1 µl of this DNA was used for PCR (20 µl reaction). Embryos from *Crt* heterozygote crosses were genotyped using two different primer sets to discriminate between wild type, heterozygous and *Crt* nulls. The *Crt* Neo3'(5'-TCGTGCTTTACGGTATCGCCGCTCCCGATT-3') and GL31 (5'-CGCGGATCCACCTCCCATGACAGCCATTTA-3') primer set detects a 600 base pair band in heterozygous embryos, and does not amplify wild type embryos, whereas the T8 (5'CTCCAGGTCCCCGTAAAATTTGCC3') and T19 (5'AGGTCTAAACCAGTCAAAAGGACC3') primer set amplifies a 715 base pair band in *Crt* WT homozygotes. The PCR conditions included initial denaturation for one cycle of 15 min at 95°C followed by 34 cycles of 30 s at 95°C, 30 s at 57°C and 30 s at 72°C.

Image collection and analysis

Mouse embryo sections and primary cultures were visualized using a Leica DMRE fluorescence microscope with 20x, 40x, and 100x objective lenses (PL Fluorotar) and Chroma filter cubes (set 31000 for DAPI, set 31002 for TRITC). Images were captured with a monochrome 10-bit CCD camera (Qimaging) and collected with Northern Eclipse v6.0 [Empix Imaging Inc.]. Images from 10T1/2 nuclei were collected using a Zeiss Axioplan 2 digital imaging microscope equipped with a 100x (1.4NA) plan-apochromat

lens and a Coolsnap HQ cooled-CCD camera (Photometrics/Roper Scientific). Images were deconvolved with Northern Eclipse v6.0 and some images were adjusted using baseline subtraction, linear stretch and brightness/contrast using Imaris 4.0 (Bitplane). Image thresholding, ratios, and linescans were performed on merged images using ImageJ and further analyzed using Microsoft Excel. In cases where linescan measurements were carried out, multiple nuclei were analyzed to establish the consistency of immunofluorescence signals with respect to DAPI. Select images were deconvolved using Huygens Essential (Scientific Volume Imaging).

Images of C2C12, RH-30, primary culture, and chick embryo nuclei were collected using a Zeiss Axioplan 2 digital imaging microscope equipped with a 100x (1.4NA) plan-apochromat lens and a Coolsnap HQ cooled-CCD camera (Photometrics/Roper Scientific). Three dimensional images for C2C12 nuclei were collected from serial images in the Z plane 0.3 μm apart. Images were deconvolved in Autoquant by blind deconvolution, Gold's algorithm for 10 iterations. For morphometic analysis, images were segmented and values from 1-255 were quantified. Volume measurements of methylation antibodies, DAPI-rich heterochromatin and total nuclear volume were collected and processed in Excel. Three dimensional images were collected from serial images in the Z plane 0.3 μm apart. Images from RH-30, primary cells, and chick embryos were analyzed in two dimensions and were deconvolved using Huygens Essential (Scientific Volume Imaging).

Histone Acid Extraction and Immunoblot Analysis

C2C12 cells were cultured in DMEM supplemented with 4 mM L-glutamine and 10% FBS to 80% confluency, induced to differentiate in reduced serum (2% sheep serum), and harvested immediately (day 1) or following 3 or 7 days of differentiation using 0.5 mM EDTA. Cells were pelleted by centrifugation at 2,000 rpm for 5 min and washed three times with 500 µl of PBS. Cells were then lysed with buffer containing 250 mM sucrose, 200 mM NaCl, 10 mM Tris-HCl (pH 8.0), 2 mM $MgCl_2$, 1 mM $CaCl_2$, 1% triton X-100, and 1 mM phenylmethanesulfonyl fluoride (PMSF). Nuclei were recovered by centrifugation at 2,000 rpm for 10 min, resuspended in 0.5 mL H_2SO_4 and left on ice for 30 min. Insoluble debris was cleared by centrifugation at 14,000 rpm for 10 min and 0.4 mL of supernatant was collected and added to 60 µL of 1M Tris (pH 8.0) and 40 µL of 10N NaOH. Equivalent quantities of protein extracts from each time point (as determined by coomassie staining and Bradford assay) were electrophoresed on 15% SDS-polyacrylamide gels and transferred to PVDF membranes. Membranes were blocked for 1 hr at RT in PBS with 5 % BSA and incubated with one of the following primary antibodies for 1 hr: rabbit-mMeK20-H4 (1/500)[AbCam], rabbit-diMeK20-H4 (1/200), rabbit-tMeK20-H4 (1/750)[AbCam], rabbit-mMeK36-H3 (1/200)[Abcam], rabbit-tMeK36-H3 (1/500) [Abcam], rabbit-mMeK79-H3 (1/500) [Abcam], rabbit-tMe79-H3 (1/500) [Abcam], and mouse-phosphoS10-H3 (1/500) [Upstate]. Immunoblots were washed 3 x 5 min in TBST and incubated with secondary antibody (anti-rabbit HRP (1/5000) or anti-mouse HRP (1/3000)) for 1 hr. Chemiluminescence was performed according to the manufacturer's

instructions (Amersham Biociences) and blots were developed on Kodak BioMax MR Film (Cat. 870-1302).

In silico analysis of histone methyltransferase gene expression

Mouse and/or Human expression data was collected for 44 SET/PR domain proteins in addition to Dot1. Using Unigene's EST profile viewer, transcripts per million for 43 HKMTases were collected in Excel [Microsoft] for 31 human tissues and assigned cutoff values relative to the mean expression of all tissues. Expression data for 43 HKMTases was obtained from SymAtlas (http://symatlas.gnf.org) and screened for elevated expression relative to the mean expression of all tissues as determined by SymAtlas. The expression of HKMTases over the course of C2C12 differentiation was obtained from a publicly available database http://www.chb-genomics.org/beggslab.

First strand cDNA synthesis and RT PCR

Total RNA was isolated from C2C12 cells using Trizol reagent (Invitrogen, as detailed by manufacturer) following 1 day of growth with 10% FBS (day 0) or 1, 3, 5 or 7 days of differentiation in 2% FBS. Contaminating DNA was eliminated by incubating samples with 1 U DNase/μg total RNA (Promega) for 15 minutes at 25 °C. The RNA was reverse transcribed using a Protoscript First Strand cDNA Synthesis Kit (New England Biolabs, as detailed by manufacturer). Primers were designed to amplify mouse Prdm12 (forward: 5'-GAGATGATCCCTCCAGACCA-3', reverse: 5'-GCACACCTGGCACTTGTAGG-3'), skm-Bop (forward: 5'-CAAAGGCAGAGGTCTGAAGG-3', reverse: 5'-TAAGTCATCCACGGACACCA-3').

The PCR conditions included initial denaturation for one cycle of 15 min at 95°C, followed by 39 cycles of 30 s at 95°C, 30 s at 60°C, and 60 s at 72°C. PCR products were analyzed by electrophoresis on a 0.5% TBE agarose gel.

DNA constructs and transfections

Constructs

Full length human *Smyd1* was subcloned into pEYFPN.1 (Clontech) from a pReceiver MO3 expression construct obtained from Genecopoeia, which contained point mutations. These mutations were repaired with a Quickchange Site-Directed Mutagenesis kit as described by the manufacturer [Stratagene] and sequence verified to generate a Smyd1-pEYFPN.1 construct with wild type sequence used in our overexpression experiments. An N-terminal flag-tagged human Prdm12-pCMVTag2A construct was obtained from S. Huang (Burnham Institute), which was truncated at the last 65 amino acids (Figure 5-4). HA-Prdm12 containing full length Prdm12 sequence was generated using PCR primers to amplify a carboxyl terminal segment of Prdm12 (350 bp) from human genomic DNA and ligated to the N-terminal sequence from the Prdm12-pCMVTag2A vector, which was cloned into a pCI-HA vector. Primers used for amplifying the Prdm12 from human genomic DNA were PRDM12F_BglII (5'-AACCTGAGATCTCACATGCGCATCCACACGCT-3') and PRDM12R_TGA KpnI (5'-CGTGGTACCTCACAGCACCATGGCCGGCA-3'), using PCR conditions described above with the exception of using a 65 °C annealing temperature and an extension time of 90 seconds. Primers used to amplify sequence from Prdm12-pCMVTag2A were

PRDM12F_PCI_HA EcoRI (5'-CAGAATTCGGCTCCGTGCTCCCGGCTGAGGCCCTGGTGCTCA-3') and PRDM12R_BglII (5'-CATGTGAGATCTCAGGTTGCTGCGCGAGTTG-3') using 65 °C for annealing temperature and an extension time of 30 seconds.

Transfections

C2C12 cells were cultured in DMEM supplemented with 4mM L-glutamine and 20% FBS to 80% confluency. RH-30 cells were cultured in RPMI media supplemented with 2mM L-glutamine and 10% FBS. Prior to transfection, both C2C12 and RH-30 cells were seeded on glass cover slips in 6-well plates at a density of ~1.0 x 10^6 cells/ml. After 18 hours of growth, cells were transfected with constructs at a 6:1 Fugene (Roche) to DNA ratio and induced to differentiate by serum withdrawal (2% FBS) 24 hrs post-transfection. Primary limb bud cultures were prepared from E10.5 embryonic limbs and cultured as previously described (Cash et al. 1997; Weston et al. 2000). Transfections in these cultures were performed by mixing 40 ul of ~2.0 x 10^6 primary cells/ml with 100 ul DMEM containing a 6:1 Fugene (Roche) to DNA and 20 ul of this mixture was placed directly on glass coveslips and incubated for 2 hrs at 37 °C + 5 % $C0_2$ for adherence. Primary culture media was then added to wells and replaced with fresh media every 24 hrs until cells were fixed for immunofluorescence as described above.

Chapter 3 ♦ Distinct distribution and dynamics of histone methyl-lysine derivatives in mouse development

* Image panels A and D in Figures 3-1, 3-2 and 3-3 were provided by Kirk J. McManus.

Versions of this chapter are published in:

1. **Vincent L. Biron**, Kirk J.McManus, Ninghe Hu, Michael J. Hendzel and D. Alan Underhill. (2004) Distinct dynamics and distribution of histone methyl-lysine derivatives in mouse development. *Dev Biol* **276 (2)**, 336-50

2. Kirk J. McManus, **Vincent L. Biron**, Ryan Heit, D. Alan Underhill and Michael J. Hendzel (2006) Dynamic Changes in Histone H3 Lysine 9 Methylations: Identification of a Mitosis-specific Function for Dynamic Methylation in Chromosome Congression and Segregation. *JBC* **281 (13),** 8888-97.

Introduction

Histone lysine methylation is known to occur on histones H3 (K4, K9, K36, K79) and H4 (K20), as mono- , di-, or trimethylated states. These modifications are catalyzed by histone lysine methyltransferases (HKMTases), which contain a conserved SET (Su-var, Enhancer of Zeste, Trithorax) domain, with the exception of Dot1. With the recent discovery of two lysine demethylases (HKDTase), specific histone demethylation reactions are now known to occur on histone H3 lysines 4, 9 and 36 (Forneris et al. 2005; Shi et al. 2005; Yamane et al. 2006). Nevertheless, lysine methylation-directed silencing of genes is thought to have the potential to convey long term epigenetic information over several cell generations during development (Margueron et al. 2005).

Several HKMTases have been shown to play essential roles in development. For instance, severe defects in embryogenesis are observed upon targeted loss of the Ezh2, G9a, NSD1, Suv39h1/Suv39h2 or Mll HKMTases (Yu et al. 1995; Peters et al. 2001; Tachibana et al. 2002; Erhardt et al. 2003; Rayasam et al. 2003). Moreover, the differences in phenotype associated with each HKMTase suggest they regulate distinct developmental events. Consistent with such an idea, studies of mono-, di-, and tri-methyl lysine 9 derivatives of histone H3 in Suv39H1/2 and G9a knockout cells reveal these HKMTases can contribute to spatially unique methylation events (Lehnertz et al. 2003). There is currently limited knowledge of how histone methyl-lysine derivatives are distributed during development and how this influences the organization of chromatin. To gain insight into this process, we have used an immunofluorescence approach to define the overall distribution of histone H3 and H4 methyl-lysine derivatives during mouse embryogenesis.

We have examined a panel of histone lysine methylations (Table 3-1) where we observed significant differences in subnuclear localization of the mono- and tri-methyl versions of histone H3 modified at K4 and K9, and of K20 on histone H4. Moreover, this differential subnuclear localization extends to distinct changes in histone H4-K20 methylation levels associated with neural and striated muscle differentiation and marked changes in tMeK9-H3 during the cell cycle (Biron et al. 2004; McManus et al. 2006). Together, these results provide novel insight into the organization of chromatin domains during embryogenesis and also suggest potential contributions of specific methyl-histone modifications to developmental processes.

Results

Immunofluorescence analysis of histone lysine methylation in mouse embryogenesis

A panel of 16 antibodies (13 methylation sites) was used to characterize histone lysine methylation during embryonic development (Table 3-1). Immunofluorescence was performed with mouse E9.5 embryos as this stage of development is marked by early stages of differentiation in many cell lineages along with a proliferative expansion of many tissues. For instance, the medial edge of the neuroepithelium undergoes rapid proliferation, whereas more lateral cells are post-mitotic and are in the early stages of differentiation (Tanabe and Jessell 1996).

Low magnification (10x) microscopic examination revealed 5 antibodies with a broad range of immunofluorescence intensity that correlated with distinct embryonic cells and/or tissues. As expected, serine 10 phosphorylated histone H3 (pS10-H3) was elevated in

proliferative regions of the embryo, supporting its previously reported association with mitotic chromosomes. Surprisingly, mMeK20-H4 and tMeK20-H4, tMeK9-H3 and tMeK9pS10-H3 were shown to have broad differences in their levels across embryonic sections. Further analysis of K20-H4 and K9-H3 methylations was undertaken to examine their spatial distributions during embryogenesis. For comparison, we also performed a detailed analysis of K4-H3 methylations as they appeared static.

Histone H3 lysine-4 methylation

Trimethylation of histone H3 on K4 (tMeK4-H3) has been shown to be associated with transcriptionally active chromatin in (Santos-Rosa et al. 2002). In contrast, dMeK4-H3 is associated with active and inactive euchromatin when examined by chromatin immunoprecipitation (ChIP) at discrete genetic loci in *S. cerevisiae* (Santos-Rosa et al. 2002). Nevertheless, how these functional associations manifest at the level of global chromatin organization is not known. Using antibodies that discriminate between the mono- and trimethyl states of K4 on histone H3, we have used immunofluorescence to characterize their localization at high resolution in 10T1/2 monolayer cells in comparison to multilayered mouse primary mesenchyme cultures and mid-gestation embryo sections. In these analyses, mMeK4-H3 was distributed in a focal pattern that was generally enriched in perinucleolar regions (surrounding perinucleolar chromatin) in 10T1/2 nuclei, as well as in mouse embryo primary cultures and sections (Figure 3-1 A-C). This association with the periphery of DAPI-rich regions was evident in linescans where the immunofluorescence peak intensities were highest immediately adjacent to those associated with the DAPI signal (*right hand* panels, Figure 3-1 A-C). In contrast, tMeK4-H3 had a qualitatively

distinct focal distribution that was distanced from the immediate edges of DAPI-dense regions in the three cell models analyzed (Figure 3-1 E-G), as apparent in linescans where the corresponding peak intensities are well separated (*right hand* panels). The tMeK4-H3 signal also encircled non-nucleolar DAPI-depleted regions and is consistent with other studies that show an accumulation of this modification in euchromatic domains adjacent to splicing speckles (Appendix, Figure A-1). Importantly, the concordance in immunofluorescence signals indicates that the global distribution of MeK4-H3 derivatives is conserved in non-immortalized cell-types (primary cultures) and within the 3D-tissue architecture of the embryo when compared with 10T1/2 nuclei. The differences in the global partitioning of mMeK4 and tMeK4-H3 indicate that even though these modifications involve the same residue, their steady-state distribution involves distinct subnuclear chromatin domains.

Histone H3 lysine-9 methylation

In contrast to the association of MeK4-H3 with euchromatin noted above, methylation of lysine 9 on the same histone has been associated with the establishment and maintenance of heterochromatin (Rea et al. 2000; Lachner et al. 2001; Peters et al. 2001). Although this relationship may be true of tMeK9-H3 (see below), mMeK9-H3 was localized to multiple foci throughout the nucleoplasmic interior that were excluded from DAPI-rich heterochromatin (Figure 3-2 A-C). Nevertheless, these foci typically abutted DAPI-rich heterochromatic regions and are therefore qualitatively similar to the foci described for mMeK4-H3 (Figure 3-1 A-C), and the two modifications produce similar linescan profiles (*right hand* panels). Consistent with the association between lysine-9 methylation and

heterochromatin formation, tMeK9-H3 co-localized extensively with DAPI-dense regions (Figure 3-2 D-F) and produced nearly identical peak intensities in linescans (*right hand* panels). In addition, tMeK9-H3 was also present in smaller foci distributed in non-heterochromatic regions of 10T1/2 nuclei (Figure 3-2 D) and to a lesser extent in nuclei from primary cultures and embryo sections (Figure 3-2 E, F). We also observed that tMeK9-H3 levels were significantly elevated in mitotic chromosomes when compared with interphase nuclei in primary cultures and embryo sections (Figure 3-2 E, F). Together with a more detailed analysis of tMeK9-H3 levels in 10T1/2 cells (McManus et al. 2006) these results suggest a relationship between this epigenetic mark and the cell cycle. Finally, as noted above for MeK4-H3, different lysine-9 methyl marks accumulate in distinct subnuclear domains and are further distinguished by differences suggestive of cell cycle influenced dynamics.

Histone H4 lysine-20 methylation

Histone H4 contains a single methyl-lysine target site at position 20, which is modified by the Pr-Set7 (Set8), NSD1 (Fang et al. 2002; Nishioka et al. 2002), Ash1 (Beisel et al. 2002) and Suv4-20h1/h2 HKMTases (Schotta et al. 2004). Pr-Set7 has been shown to be a K20-H4 monomethylase, whereas the Suv4-20h HKMTases appear to preferentially catalyze tMeK20-H4 (Schotta et al. 2004; Xiao et al. 2005), but their roles in development are not known. In 10T1/2 cells, mMeK20-H4 was distributed in a focal pattern in nucleoplasmic regions excluded from nucleoli and DAPI-dense heterochromatin (Figure 3-3 A). Further analysis using linescans show that high intensity DAPI and mMeK20-H4 peaks do not coincide, but are juxtaposed, indicating that

mMeK20-H4 is enriched at the periphery of heterochromatic regions. In addition, a small percentage of cells (~10-15%) displayed large non-heterochromatic subnuclear accumulations of mMeK20-H4 which were more prominent in muscle cell cultures (~>95%) (detailed in Chapter 4). Although the focal distribution was conserved in interphase nuclei, including embryonic primary cells and sections (Figure 3-3 A, B), we observed significant variations in signal intensity between nuclei that suggested cell cycle associated changes. Quantitative measurements in 10T1/2 nuclei (McManus et al. 2006) demonstrate significant variation in global mMeK20-H4 levels that are cell cycle associated. These results suggest that mMeK20-H4 is found in chromatin regions that are non-repressive in nature and that global mMeK20-H4 levels are regulated during interphase and mitosis. Characterization of tMeK20-H4 distribution revealed its accumulation primarily in DAPI-dense regions of 10T1/2 nuclei, which are characteristic of pericentromeric heterochromatin, and in additional small foci dispersed throughout intermediate DAPI staining regions (Figure 3-3 D). This distribution is consistent with recent studies reporting the subnuclear distribution of tMeK20-H4 in mammalian cells, using independently generated antibodies (Kourmouli et al. 2004; Schotta et al. 2004). In limb mesenchyme cultures and embryo sections, however, tMeK20-H4 was exclusively co-localized with DAPI-dense regions (Figure 3-3 E, F) and was therefore more restricted than in 10T1/2 cells. Accordingly, linescans detected essentially no signal intensity outside of these centromeric domains (compare *right hand* panels in Figure 3-3 E, F to D). These results for lysine 20 again illustrate that the two methylation states partition to

different regions of the nucleus, and reveal potential cell cycle-associated changes in the mono-methyl modification of this residue.

Regional differences in methyl histone derivatives

The above analyses reveal several important trends in the distribution of methyl-lysine modifications on histones H3 and H4. First, in each case, the mono- and tri-methyl derivatives accumulate in distinct nuclear regions, which strongly suggest they are associated with distinct chromatin environments. Second, the mMeK9-H3 and mMeK20-H4 modifications appear to exhibit cell cycle dependent differences. Adding to these observations, we also detected significant regional differences for histone lysine methylation in the developing embryo and during the course of cellular differentiation. As an example, the distribution in the neural tube at E9.5 is described because it contains spatially distinct populations of proliferating and differentiating neural cells. For these analyses, merged images of DAPI and either the mono- or tri-methyl modifications were examined after controlling for regional increases in cellularity by calculating a ratio image of TRITC/DAPI. Furthermore, the intensity distributions of mono- and tri-methyl derivatives were displayed as high and low intensity values using separate indexed colors to expose any intensity gradients throughout the neural tube. For mMeK4-H3, tMeK4-H3 and mMeK9-H3, the local increases observed in the neural tube are due to elevated cellularity, and high intensity values are randomly distributed (Figure 3-4 A-C), providing an important control for the remaining epigenetic marks. Furthermore, these modifications appeared randomly distributed throughout all other developing tissues in mid-gestation embryos (data not shown). In contrast, tMeK9-H3 was characterized by discrete

fluorescent signals in a portion of the cells at the medial edges of neural tube. Notably, mitotic nuclei lining the lumen of the neural tube (Figure 3-4 D) and the periphery of somites (not shown) in E9.5 embryo sections were significantly elevated above other nuclei that were in interphase. Similar tissue distributions were observed for pS10-H3 and pS10tMeK9-H3 (data not shown). In this regard, the specific elevation of tMeK9-H3 levels in mitotic nuclei found in the embryo provides a useful and novel marker to define regions of the embryo that are rapidly proliferating.

We also observed significant disparity in the distribution of mMeK20-H4 and tMeK20-H4 in mid-gestation mouse embryos. At E9.5, mMeK20-H4 formed a medial to lateral gradient where the luminal cells displayed the highest signal intensity (Figure 3-4 E), a region that is populated by rapidly dividing neuroblasts that migrate laterally and subsequently undergo neuronal differentiation (Jessell 2000; Marquardt and Pfaff 2001). This is somewhat different from the elevated pattern noted for tMeK9-H3 in the medial neural tube, which was restricted to mitotic cells along the neural tube lumen. In contrast to mMeK20-H4 and tMeK9-H3, tMeK20-H4 levels were reduced in this portion of the neural tube and elevated in lateral regions at E9.5 (Figure 3-4 F). Taken together, these results demonstrate that the degree of histone lysine methylation spatially and temporally correlates with distinct cellular events during development.

Spatiotemporal distributions of histone H4 lysine-20 methyl derivatives

Neurogenesis

We examined the neuroepithelial distributions of histone H4-K20 methylation in further detail. Although mMeK20-H4 retained its association with the medial portion of the neural

tube at all axial levels and developmental stages examined (Figure 3-4 E, Figure 3-5 A-D), tMeK20-H4 became increasingly restricted to the ventrolateral aspects by E9.5 (Figure 3-4 F, Figure 3-5 E-H) and later developmental stages (E11.5, E12.5 data not shown) after displaying a uniform distribution throughout the neuroepithelium at E8.5 (Figure 3-5 E). In fact, tMeK20-H4 immunofluorescence was virtually undetectable within nuclei in medial regions of the neural tube by E9.5 (Figure 3-5 F). Notably, this distribution of tMeK20-H4 in the neuroepithelium coincides with the development of post-mitotic spinal motor neurons, which begin to populate the ventrolateral neuroepithelium at E9-9.5 and mature to completion by E10.5-E11, and with the emergence of dorsal root ganglia (DRG) at E9.5 (Arber et al. 1999). Using chick embryos, we examined the co-localization between tMeK20-H4 and ventrolateral neuronal markers Isl1 and MNR2. We observed partial overlap (data not shown), indicating that tMeK20-H4 enrichment is associated with more than one neural subtype in the ventrolateral neuroepithelium. Taken together our results indicate that during neurulation, mMeK20-H4 is elevated in proliferating neural cells, whereas the tMeK20-H4 modification dramatically declines in these cells and increases in differentiating neurons.

To investigate the general behavior of MeK20-H4 during neural cell differentiation in further detail, we examined its distributions during retinogenesis. The chick retina is morphologically and molecularly well defined, and offers an excellent model for neuronal cell differentiation (Witte and Godbout 2002). Chick retinal neurogenesis commences at embryonic day 2 and is complete by day 13, generating six neuronal and one glial cell type. Retinogenesis is initiated by differentiation of retinal ganglion and horizontal cells,

followed by cone-photoreceptors, amacrine cells, rod-photoreceptors, bipolar cells, and müller glia (Marquardt and Gruss 2002). The subnuclear distributions of mMeK20-H4 and tMeK20-H4 in chick eye sections recapitulated the pattern observed in mouse embryo sections and C2C12 skeletal muscle cells (Figures 3-6 A, 3-7A). In addition, striking dynamics were observed in their overall levels during the course of chick retinogenesis. As expected, elevated mMeK20-H4 levels were associated with proliferative zones of the retina, as shown by association with the S-phase marker PCNA (Figure 3-6). In later stages of retinal development when proliferation levels have decreased, elevated levels of mMeK20-H4 are found in very few nuclei, consistent with a significant decrease in overall PCNA staining. Significant changes in the distribution of tMeK20-H4 were observed over the course of retinogenesis (Figure 3-7). In the retina of day 5 embryos, which contain numerous proliferative cells, an enrichment of tMeK20-H4 was found in a broad population of nuclei. In day 7 retina, however, elevation of tMeK20-H4 was restricted to cells near the nerve fiber layer, in addition to a few cells within the neural retina. By day 16, elevated of tMeK20-H4 were restricted to the ganglion cell layer. Because cells adjacent to the ganglion cell layer are also in later stages of neural differentiation, this result suggests that elevation of tMeK20-H4 is not solely dependent upon general differentiation, and appears to have a more cell type specific role in neurogenesis.

Striated muscle

Elevations of tMeK20-H4 were also apparent in non-neural tissues of the embryo at E9.5 and E10.5, notably in the heart and medial portion of somites. We therefore investigated the possibility that histone H4-K20 methylation levels may be differentially regulated in

developing striated muscle, a cell type common to both tissues. To address this, we carried out co-immunofluorescence analyses with either mMeK20-H4 or tMeK20-H4 and an antibody that detects myosin heavy chain (MF20) in skeletal and cardiac muscle. At E10.5, MF20-positive cells of the myotome were depleted in mMeK20-H4 (Figure 3-8 E) and enriched in tMeK20-H4 (Figure 3-8 F). The distribution of tMeK20-H4 was initially broad (E 8.5) and became progressively restricted to differentiated cells, whereas the depletion in mMeK20-H4 was seen from E8.5 onward but became more pronounced with an increase in differentiated cells (Figure 3-8). This inverse relationship was also found in cardiac musculature where mMeK20-H4 levels were relatively low or absent in MF20-positive cardiomyocytes (Figure 3-9 A,C,E), while tMeK20-H4 levels were progressively elevated in MF20 positive differentiating cardiomyocytes (Figure 3-9 B,D,F). These results indicate that histone H4-K20 methylation undergoes significant changes during the differentiation of striated muscle.

K20-H4 methylation in caltreticulin deficient embryos

We investigated whether tMeK20-H4 enrichment in cardiac muscle would persist in embryos with perturbed development of this tissue. For this analysis we used Calreticulin deficient mice, which have cardiac malformations that are embryonic lethal (Michalak et al. 2004). In comparison to wild type littermates, E10.5 calreticulin mutant embryos had an underdeveloped heart and limbs on gross appearance. We observed a loss of tMeK20-H4 enrichment in the cardiac cells in Calreticulin deficient embryos, where its levels were comparatively higher outside of the cardiac field (Figure 3-10). This evidence argues against a non-specific temporal accumulation of tMeK20-H4, as the loss of tMeK20-H4

was not observed in a time-matched littermate. This result also suggests an important role of tMeK20-H4 enrichment in cardiac development.

Antibody	Genome wide enrichment	Enrichments in embryos?
H3 acetylated K4+K9	euchromatin	no
H3 monoMeK4	euchromatin	no
H3 diMeK4	euchromatin	no
H3 triMeK4	euchromatin	no
H3 monoMeK9	euchromatin	no
H3 triMeK9	constitutive heterochromatin	yes
H3 pS10+triMeK9	constitutive heterochromatin	yes
H3 pS10	constitutive heterochromatin	yes
H3 diMeK23	euchromatin	no
H3 monoMeK36	euchromatin	no
H3 triMeK36	euchromatin	no
H3 monoMeK79	euchromatin	no
H3 diMeK79	euchromatin	no
H3 triMeK79	euchromatin	no
H4 monoMeK20	euchromatin and facultative heterochromatin	yes
H4 diMeK20	euchromatin	no
H4 triMeK20	constitutive heterochromatin	yes

Table 3- 1: Summary of histone post-translational modifications examined in mouse mid-gestation embryo sections.

Figure 3- 1: Subnuclear distribution of histone H3 mono- and tri-methyl lysine 4 in mouse immortalized, primary and embryo cells.

Immunofluorescence of (A-C) mMeK4-H3 and (D-F) tMeK4-H3 (TRITC column) in (A, C) 10T1/2 cells, (B, E) day 1 mouse primary cultures from E 10.5 limb buds, and (C, F) E8.5 CD1 embryo transverse sections (~mid-axial level) with respect to DAPI. Merged images were pseudo-colored (DAPI: green, TRITC: red) and analyzed for co-localization by linescan measurements through merged images (red line: TRITC, green line: DAPI). Scale bar, 3 μm. The linescan measurements are representative and were consistent over the analysis of multiple nuclei.

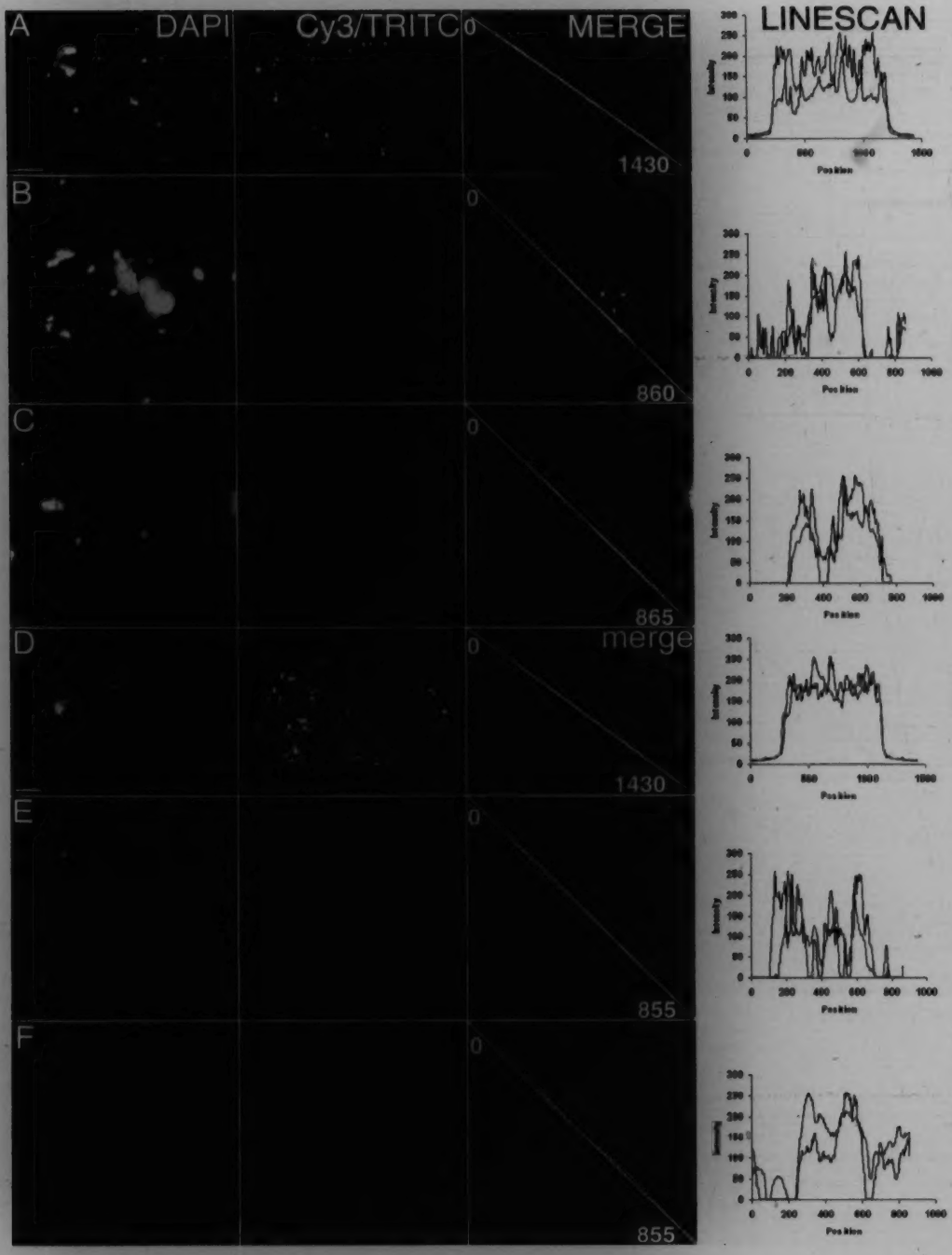

A
DAPI
Cy3/TRITC
MERGE
LINESCAN
1430
B
860
C
865
D
merge
1430
E
855
F
855

Figure 3- 2: Subnuclear distribution of histone H3 mono- and tri-methyl lysine 9 in mouse immortalized, primary and embryo cells.

Immunofluorescence of (A-C) mMeK9-H3 and (D-F) tMeK9-H3 (TRITC column) in (A, C) 10T1/2 cells, (B, E) day 1 mouse primary cultures from E 10.5 limb buds, and (C, F) E8.5 CD1 embryo transverse sections (~ mid-axial level) with respect to DAPI. Merged images were pseudo-colored (DAPI: green, TRITC: red) and analyzed for co-localization by linescan measurements through merged images (red line: TRITC, green line: DAPI). Scale bar, 3 μm. The linescan measurements are representative and were consistent over the analysis of multiple nuclei.

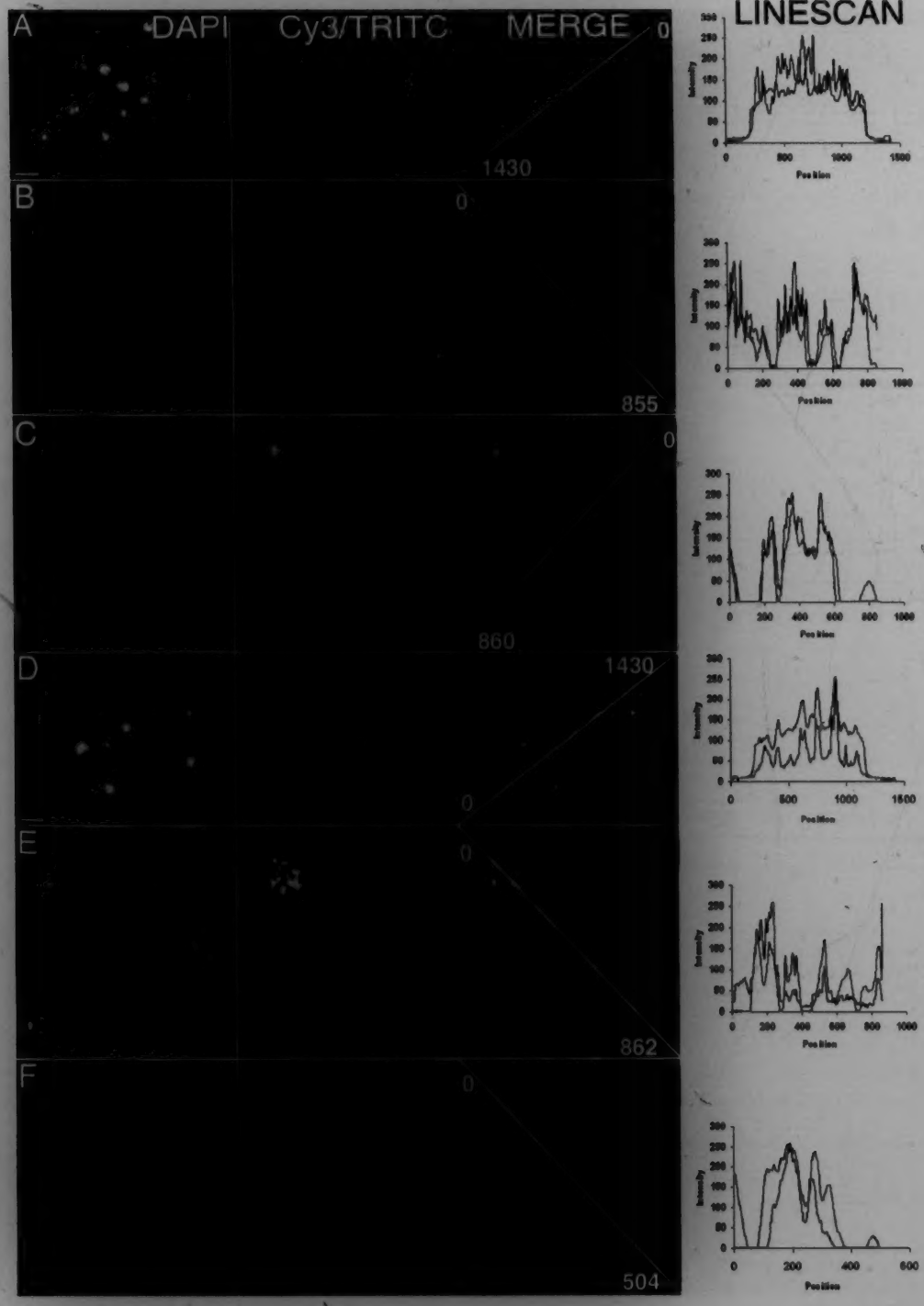

A
DAPI
Cy3/TRITC
MERGE
0
1430
B
0
855
C
0
860
D
1430
0
E
0
862
F
0
504
LINESCAN

Figure 3- 3: Subnuclear distribution of histone H4 mono- and tri-methyl lysine 20 in mouse immortalized, primary and embryo cells.

Immunofluorescence of (A-C) mMeK20-H4 and (D-F) tMeK20-H4 (TRITC column) in (A, C) 10T1/2 cells, (B, E) day 1 mouse primary cultures from E 10.5 limb buds, and (C, F) E8.5 CD1 embryo transverse sections (~mid-axial level) with respect to DAPI. Merged images were pseudo-colored (DAPI: green, TRITC: red) and analyzed for co-localization by linescan measurements through merged images (red line: TRITC, green line: DAPI). Scale bar, 3 μm. The linescan measurements are representative and were consistent over the analysis of multiple nuclei.

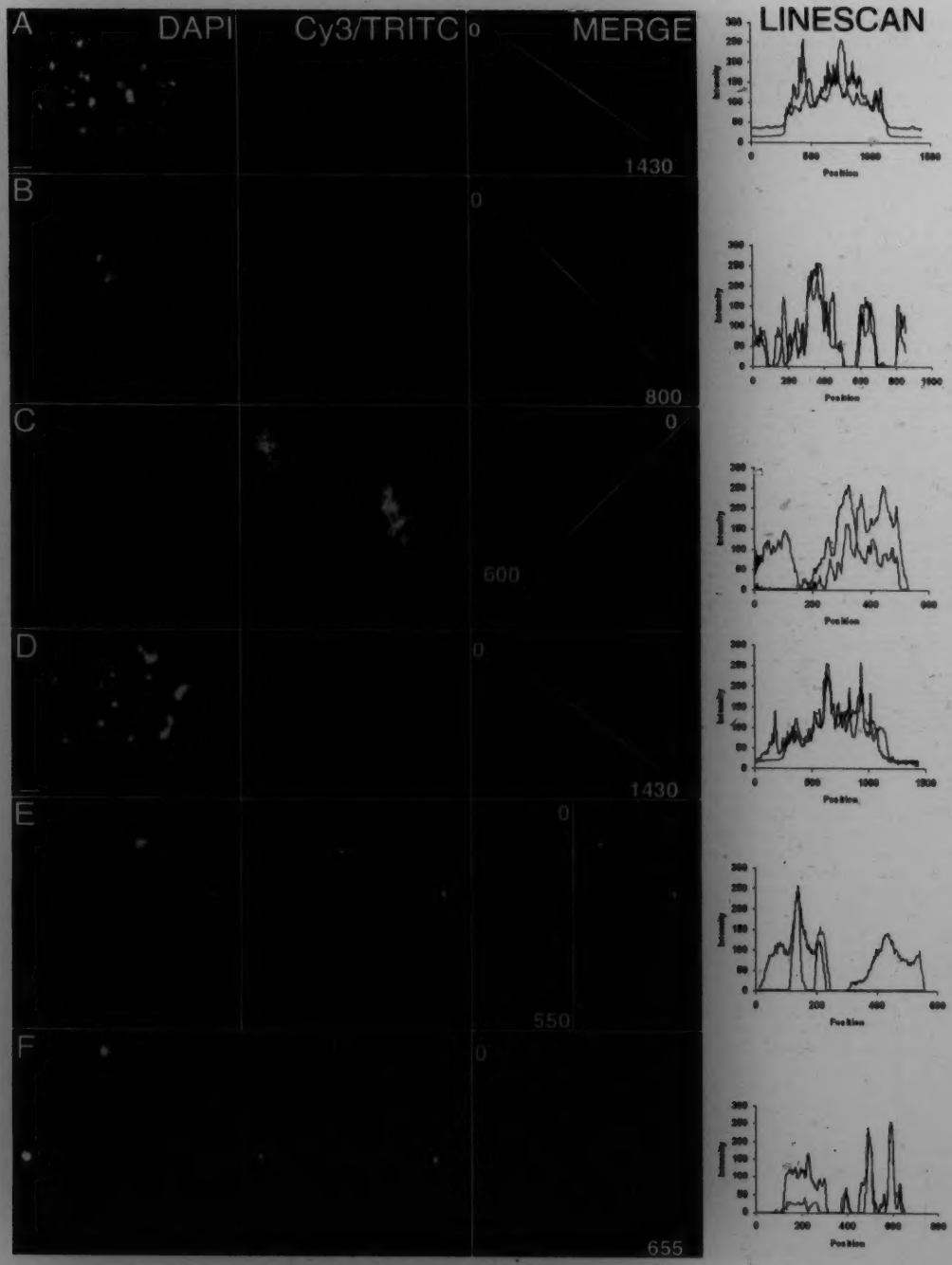

A
DAPI
Cy3/TRITC
MERGE
0
1430
LINESCAN
B
0
800
C
0
600
D
0
1430
E
0
550
F
0
655

Figure 3- 4: Histone methyl lysine distribution in mouse neuroepithelium.

Neural tube distribution of (A) mMeK9-H3, (B) H4K9me3, (C) mMeK20-H4, and (D) tMeK20-H4 in transverse sections (~ mid-axial level) of E 9.5 mouse embryo. Left column panels show histone antibodies (red) and with respect to DAPI as merged images (center column, DAPI:green, TRITC:red). DAPI stain in these panels is added to demonstrate the distribution of nuclei throughout the neuroepithelium, and the ratio of DAPI to antibody is shown in center panels. The right hand panels display intensity distributions of each methylation for cells that produce immunofluorescence signals below (blue; lower 90^{th} percentile) and above (red; upper 10^{th} percentile) a threshold defined by intensity distribution graphs. M and L indicate medial and lateral aspects of the neuroepithelium respectively.

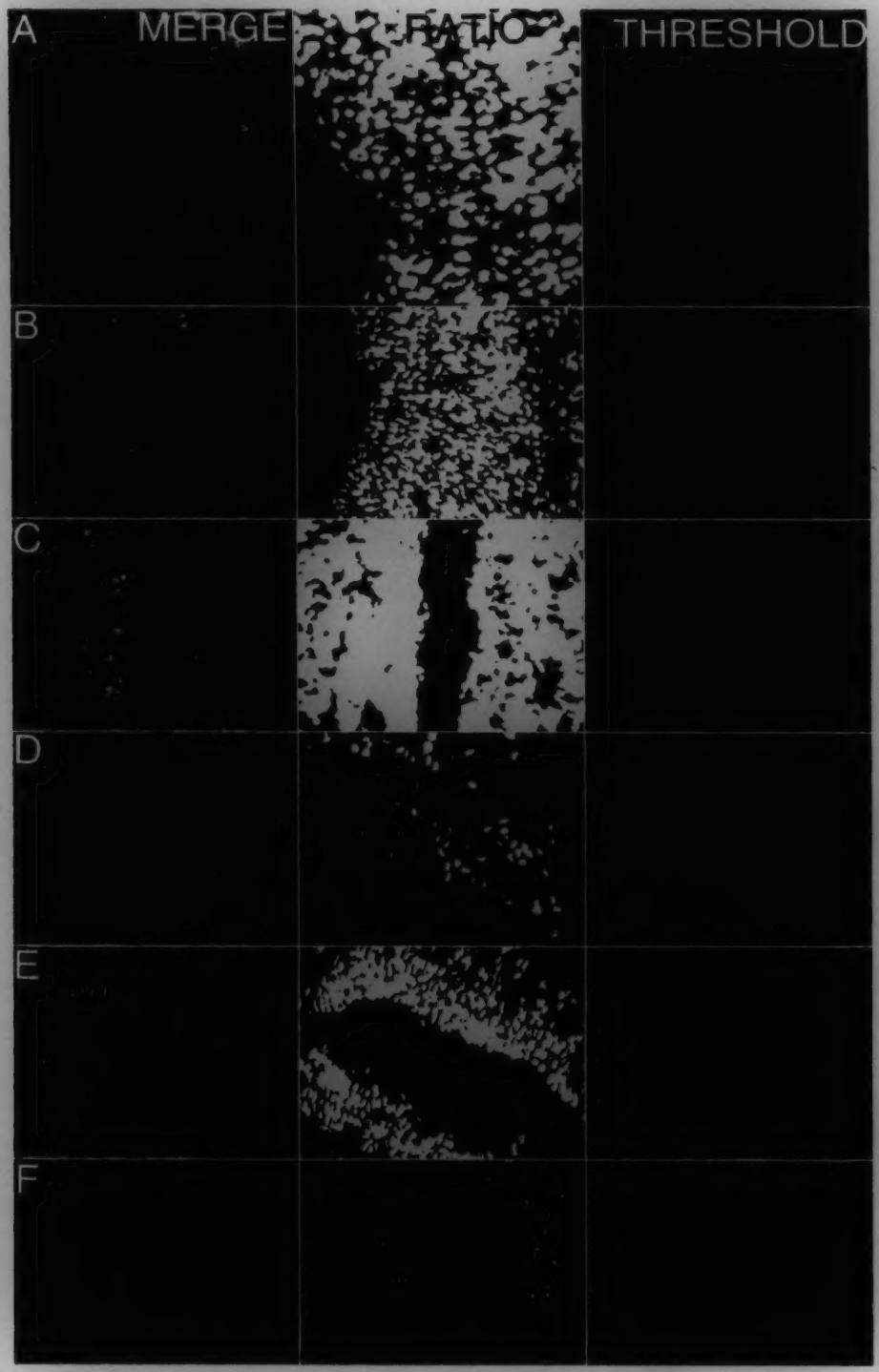

A
MERGE
RATIO
THRESHOLD
B
C
D
E
F

Figure 3- 5: Spatiotemporal dynamics of H4 mono- and tri-methyl lysine 20 in mouse neural development.

Immunofluorescence of mMeK20-H4 in developing neural tube of mouse embryo transverse sections at (A) E8.5, (B) E9.5, (C, D) E10.5 and of tMeK20-H4 at (E) E8.5, (F) E9.5 and (G, H) E10.5. D and H are sections through the hindbrain to illustrate the conserved distribution of these antibodies in the cranial neural tube and other sections are at mid-axial levels. The antibody distribution is shown with respect to nuclei by DAPI staining for both mMeK20-H4 (A'-D', DAPI=green, mMeK20=red) and tMeK20-H4 (E'-H', DAPI=green, tMeK20=red). DRG, dorsal root ganglion.

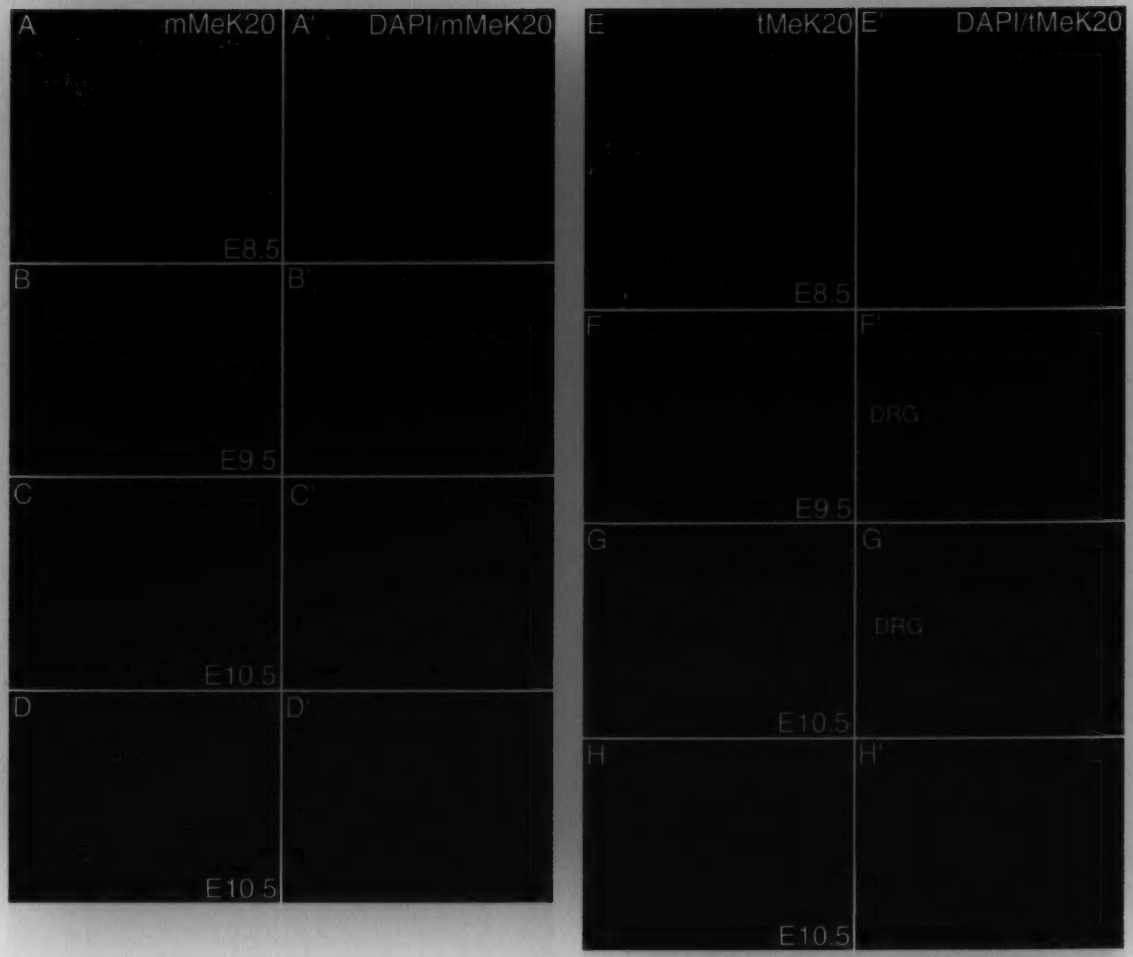

A
mMeK20
A'
DAPI/mMeK20
E8.5
B
B'
E9.5
C
C'
E10.5
D
D'
E10.5
E
tMeK20
E'
DAPI/tMeK20
E8.5
F
F'
DRG
E9.5
G
G'
DRG
E10.5
H
H'
E10.5

Figure 3- 6: Dynamics of mMeK20-H4 in developing chick retina.

A) Distribution of mMeK20-H4 in chick retinal cell nuclei with respect to DAPI. (B-D) Immunolocalization of mMeK20-H4 (red) with respect to PCNA (green) in eye sections of B) D5, C) D7 and D) D16 chick embryos (merged in rightmost panels). DAPI is shown in grayscale in leftmost panels for identification of retinal cell layers. RPE, retinal pigment epithelium; NR, neural retina; NFL, nerve fiber layer; ONL, outer nuclear layer; INL, inner nuclear layer; IPL, inner plexiform layer; GCL, ganglion cell layer.

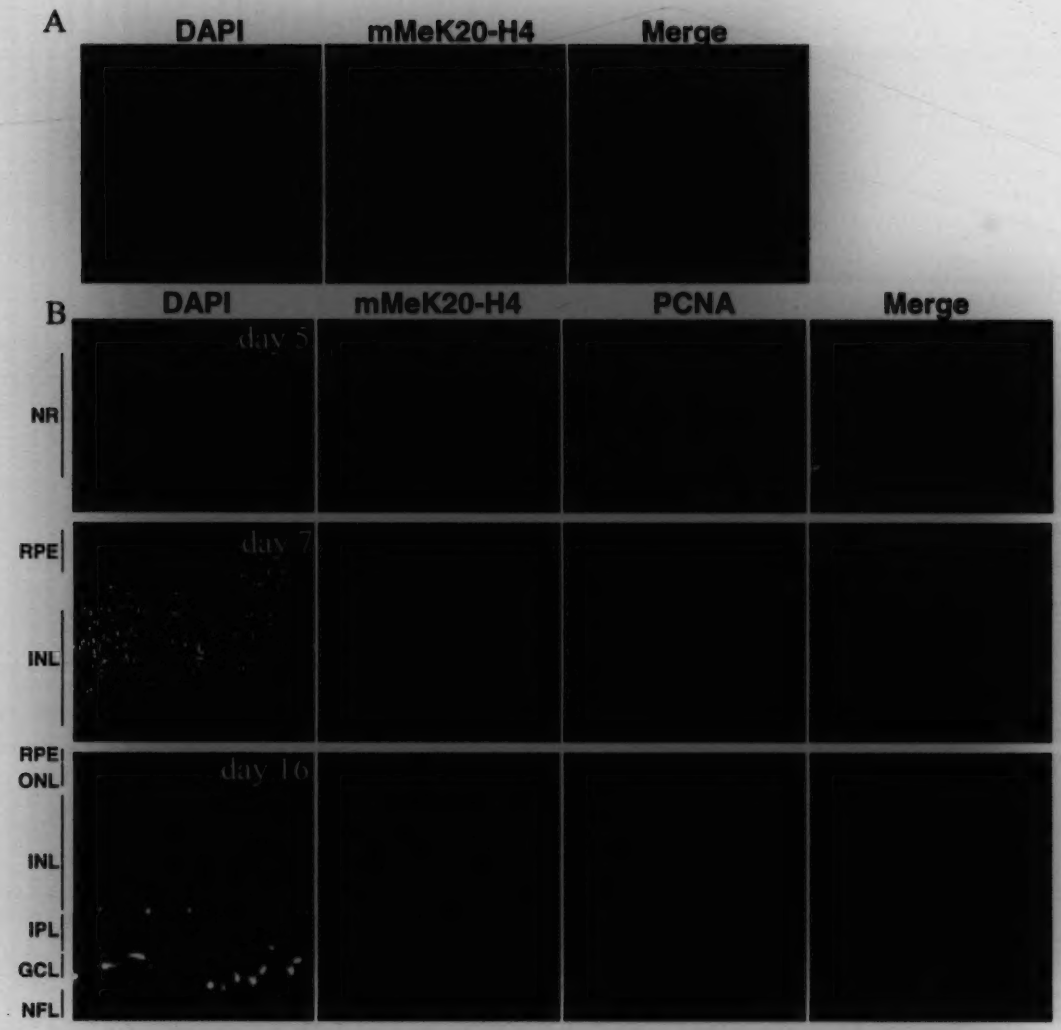

A
DAPI
mMeK20-H4
Merge
B
DAPI
mMeK20-H4
PCNA
Merge
day 5
NR
day 7
RPE
INL
day 16
RPE
ONL
INL
IPL
GCL
NFL

Figure 3- 7: Dynamics of tMeK20-H4 in developing chick retina.

(A) Distribution of tMeK20-H4 in chick retinal cell nuclei with respect to DAPI. (B-D) Immunolocalization of tMeK20-H4 (red) in eye sections of B) D5, C) D7 and D) D16 chick embryos, merged with respect to DAPI (blue). DAPI is shown in grayscale in leftmost panels for identification of retinal cell layers. RPE, retinal pigment epithelium; NR, neural retina; NFL, nerve fiber layer; ONL, outer nuclear layer; OPL, outer plexiform layer; INL, inner nuclear layer; IPL, inner plexiform layer; GCL, ganglion cell layer.

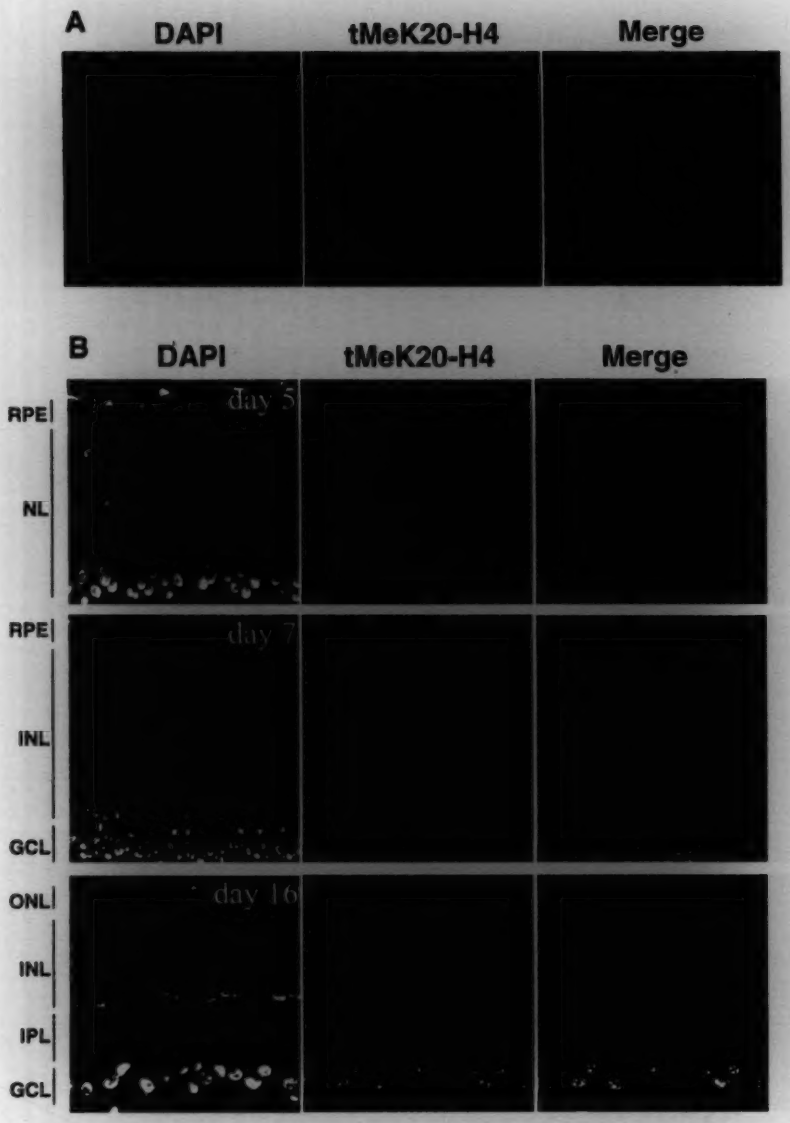

A
DAPI
tMeK20-H4
Merge
B
DAPI
tMeK20-H4
Merge
RPE
NL
day 5
RPE
INL
GCL
day 7
ONL
INL
IPL
GCL
day 16

Figure 3- 8: Differential distributions of mMeK20-H4 and tMeK20-H4 in mouse myotome development.

Immunofluorescence of mMeK20-H4 at (A) E8.5, (C) E9.5 and (E) E10.5 and (B) tMeK20-H4 at E8.5, (D) E9.5 and (F) E10.5 in transverse sections of developing somites (center panels). H4-K20 antibodies are shown with respect to DAPI to stain for nuclei (A, B; left panels), and MF-20 to detect myosin heavy chain in the myotome (C-F, left panels). MF-20 signal was not seen until E9.5. Merged images from left and center panels are shown in right panels (DAPI=blue, m/tMeK20=red, MF-20=green). NT, neural tube; My, myotome.

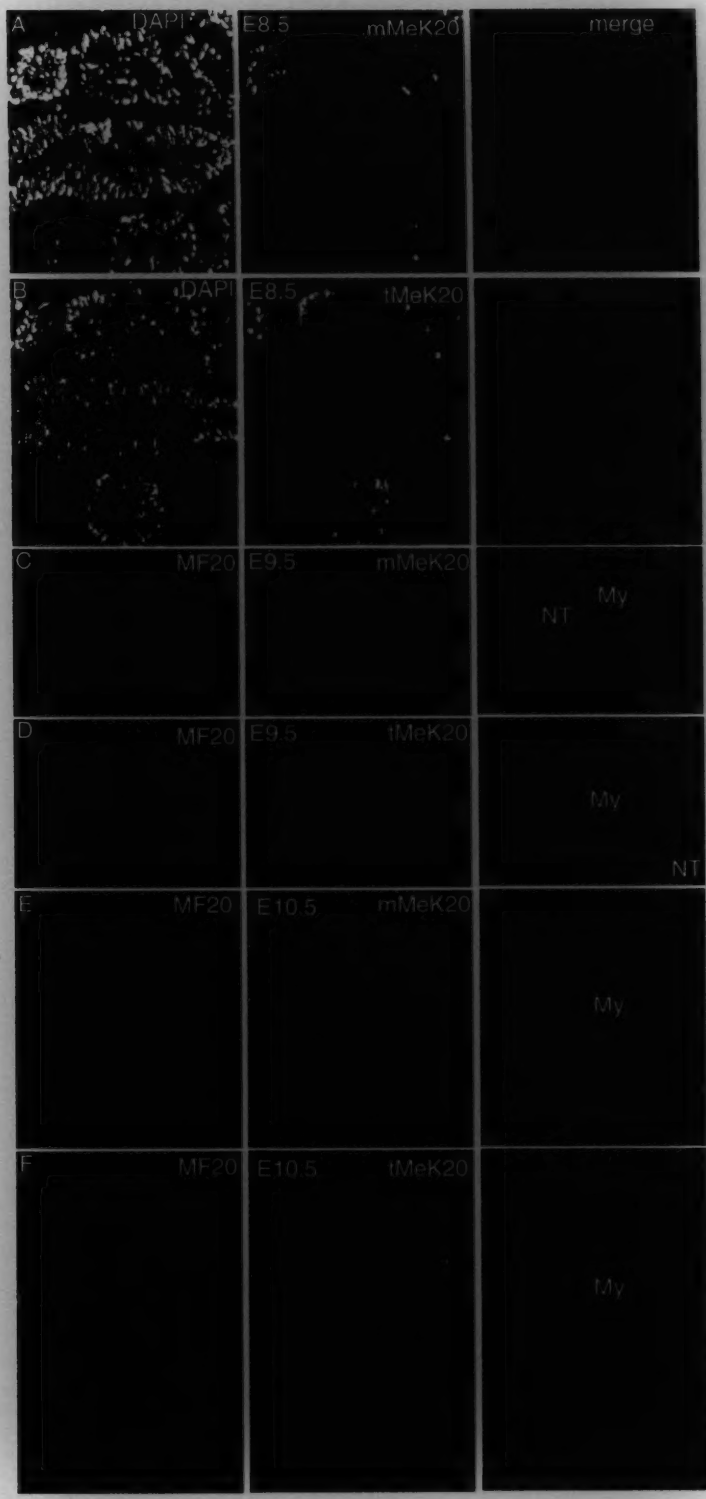

A
DAPI
E8.5
mMeK20
merge
B
DAPI
E8.5
tMeK20
C
MF20
E9.5
mMeK20
My
NT
D
MF20
E9.5
tMeK20
My
NT
E
MF20
E10.5
mMeK20
My
F
MF20
E10.5
tMeK20
My

Figure 3- 9: Differential distributions of H4 mono- and trimethyl lysine 20 in mouse cardiac muscle development.

Immunofluorescence of mMeK20-H4 at (A) E8.5, (C) E9.5 and (E) E10.5 and tMeK20-H4 at (B) E8.5, (D) E9.5 and (F) E10.5, mouse sections showing the developing heart (left panels). Cardiac muscle is shown by myosin heavy chain immunofluorescence using an MF-20 antibody (center panels). Left and center panels are merged with DAPI to illustrate cellularity (right panels; m/tMeK20=red, MF20=green, DAPI=blue). N, nuclear staining of tMeK20; C, cytoplasmic background staining.

A
mMeK20
MF20
DAPI/MF20/mMeK20
E8.5
B
tMeK20
DAPI/MF20/tMeK20
E8.5
C
mMeK20
DAPI/MF20/mMeK20
E9.5
D
tMeK20
DAPI/MF20/tMeK20
E9.5
E
mMeK20
DAPI/MF20/mMeK20
E10.5
F
tMeK20
DAPI/MF20/tMeK20
E10.5

Figure 3- 10: Disrupted levels of histone H4 trimethyl lysine 20 in cardiac muscle of calreticulin null mice.

Immunofluorescence of tMeK20-H4 (left panels) in cardiac muscle, detected by myosin heavy chain (MF-20, center panels), of a *Crt* null embryo at E10.5, with respect to DAPI (merged right panels) (DAPI=blue, tMeK20-H4=red and MF-20=green). Note the disrupted morphology of *Crt* -/- hearts in compassion to a wild type (WT), which is sectioned transversely approximately mid way through the heart. It should also be noted that the *Crt* -/- heart is much smaller and comparable to that of an E 8-9 embryo.

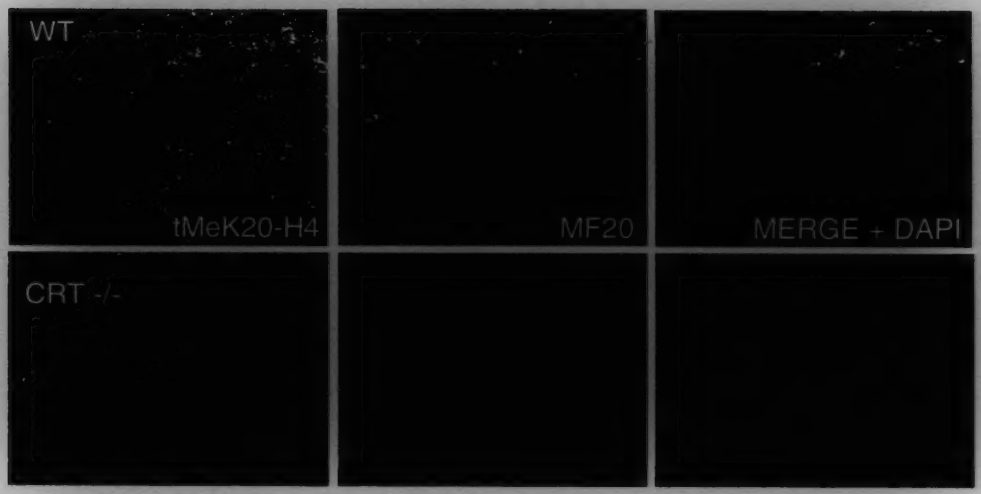
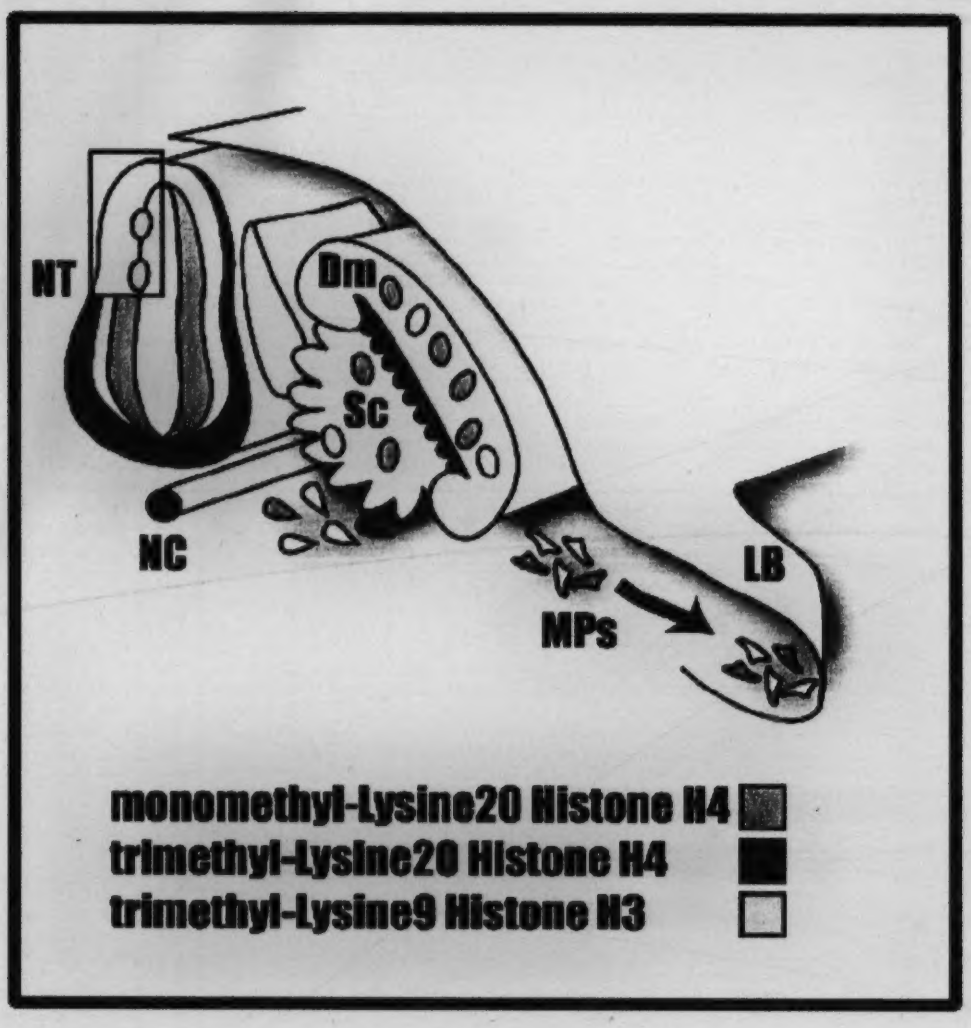

WT
tMeK20-H4
MF20
MERGE + DAPI
CRT -/-

Figure 3- 11: Summary for histone methylation in mouse embryonic development

During mouse mid-gestation, mMeK20-H4 (blue) levels are elevated in proliferative cells found in higher concentration in regions such as the medial neural tube and the somite periphery. These elevated levels are present in both interphase and mitotic nuclei, in contrast to tMeK9-H3 (orange, inset), which is highly elevated in mitotic nuclei, found within mMeK20-H4 elevated cell populations. In contrast, tMeK20-H4 levels are progressively enriched in differentiating neurons found in lateral aspects of the neural tube and skeletal muscle cells concentrated in the myotome. These marked differences in mMeK20-H4 and tMeK20-H4 dynamics may be epigenetic determinants of cellular proliferation and differentiation required for mouse development. NT, neural tube; NC, notochord; Dm, dermomyotome; Sc, sclerotome; MPs, myogenic precursors; LB, limb bud.


NT
Dm
Sc
NC
MPs
LB
monomethyl-Lysine20 Histone H4
trimethyl-Lysine20 Histone H4
trimethyl-Lysine9 Histone H3

Discussion

Histone lysine methylation is known to occur on the amino-terminal tails of histone H3 at lysines 4, 9, 27 and 36, and of lysine 20 in histone H4 (Lachner et al. 2003), in addition to new modifications identified by mass spectrometry (Cocklin and Wang 2003). Moreover, the ability of lysine residues to acquire mono-, di-, and tri-methyl moieties has the potential to generate extensive functional diversity through the recruitment of distinct regulatory proteins or complexes (Fischle et al. 2003). Despite recent findings, the biological relevance of a histone methyl-lysine hierarchy remains poorly understood and the spatial distribution has not been defined for the majority of methyl states. To address this issue, we have used a developmental model system to examine specific methyl-histone derivatives in the context of cellular proliferation, specification, and terminal differentiation. Based on the results of this analysis, we suggest that the degree of lysine methylation, as well as the residue modified, demarcates distinct and dynamic chromatin environments at both the cellular and tissue level.

Traditionally, chromatin has been categorized as highly condensed heterochromatin or less compact euchromatin, where the latter is considered the template for gene expression (Weintraub and Groudine 1976; Vermaak et al. 2003). Based on this two-state model, particular histone modifications should segregate to either of these chromatin domains. The results presented here, which describe multiple, distinct steady-state distributions for histone methyl-lysine derivatives (Figures 3-1 – 3-3), suggest additional levels of organization exist. Within this scheme, the mono-methyl modifications displayed a similar distribution to punctate foci adjacent to regions of heterochromatin, while the tri-

methyl versions occupied chromatin domains with clear functional significance. For tMeK9-H3 and tMeK20-H4, this involves pericentric heterochromatin, while tMeK4-H3 localizes to the euchromatin compartment and has recently been shown to be a specific epigenetic mark for active chromatin in *S. cerevisiae* (Santos-Rosa et al. 2002) and higher eukaryotes (Schneider et al. 2004). In a separate study, we have established that tMeK4-H3 associate with the periphery of splicing speckle (see Appendix Figure A-1), complementing the previous definition of speckle-associated chromatin domain that accumulates hyperacetylated forms of histone H3 (Hendzel et al. 1998). Taken together, the tri-methyl modified forms therefore appear to confer a greater influence on chromatin structure and function – or 'a more robust epigenetic imprint' as suggested by Peters et al. (Lehnertz et al. 2003) – than the mono-methyl lysine modified histones.

The markedly distinct locations of the mono- and tri-methyl forms can be accommodated in two ways. The first scenario involves a 'static' model where the conversion of mono-methyl to tri-methyl modified lysine residues occurs within a spatially fixed chromatin domain that is separate from the mono-methyl domains. An expectation of this model is that the activity of underlying HKMTase(s) is confined, which is observed for the Suv39h1 and Suv39h2 HKMTases and their placement of the tMeK9-H3 mark on pericentric heterochromatin (Lehnertz et al. 2003). Notably, in the absence of both *Suv39h1* and *Suv39h2*, the loss of tMeK9-H3 is confined to pericentromeric heterochromatin where there is also a concomitant increase in mMeK9-H3 (Lehnertz et al. 2003). Although this 'static' model can account for some aspects of the steady state distributions observed for methyl-modified histones, chromatin also exhibits both

constrained motion and long-range movement that may also influence these patterns (Marshall et al. 1997; Belmont 2001; Tumbar and Belmont 2001; Vazquez et al. 2001; Marshall 2002; Belmont 2003; Janicki and Spector 2003). For instance, the α and β-globin genes relocate within the nucleus in response to their activation or repression status (Brown et al. 2001), and recombinant transgenes integrated into the same locus relocate to centromeric heterochromatin when silenced (Francastel et al. 1999). Similarly, recruitment of HP1α to an active euchromatic locus has been shown to cause histone K9-H3 methylation and moreover, relocation of the locus to heterochromatic regions that together create a repressed chromatin state (Ayyanathan et al. 2003). In this 'dynamic' model, the initial localization of the mono-methyl derivatives may constitute a staging area where the conversion to the tri-methyl form is accompanied by changes in chromatin localization within the nucleus. Importantly, this could account for the marked changes in mMeK20-H4 and tMeK20-H4 observed in embryogenesis where global changes in gene activity during cell differentiation would be expected to cause changes in the levels and distributions of methyl-lysine derivatives. Thus, the differences in distribution of methyl-modified histones noted here suggest they have the potential to define regional chromatin specialization within the nucleus.

Dynamics of histone lysine methylation in development

A key element of the model system used here is its access to a broad range of cell-types at various levels of commitment and differentiation. Within this paradigm, the behavior of tMeK9-H3, mMeK20-H4, and tMeK20-H4 has provided novel insight into methyl-histone biology and its potential roles in development. In the case of tMeK9-H3, this was

revealed by its marked elevation in mitotic cells and its absence in interphase cells lining the neural tube lumen, which is underscored in comparison with the uniform labeling of the neuroepithelium with mMeK9-H3 (Figure 3-4 C and D). Given this distribution of tMeK9-H3, we suggest it constitutes a novel and specific epigenetic mark for cell proliferation in development. This behavior also reveals an unforeseen dynamic range for a methyl histone modification that is consistent with changes in tMeK9-H3 in cell culture (McManus et al. 2006). In such experiments, the analysis of tMeK9-H3 levels in synchronized populations by flow cytometry, immunoblot analysis, and quantitative imaging clearly indicates a loss of signal intensity as cells leave mitosis. Moreover, this decline is delineated from replication-associated histone deposition or transcription-coupled histone replacement, and suggests some other mechanism must account for removal of the tMeK9-H3. In this regard, the dramatic, genome-wide loss of tMeK9-H3 raises the possibility of its removal by a histone demethylase. Recent data that was not available at the time this research project was undertaken, suggests the possibility of such an enzyme. The flavin-dependent amine oxidase LSD1 was the first discovered histone lysine demethylase, shown to specifically demethylate H3 mono- and diMeK4 or H3 diMeK9 (Forneris et al. 2005; Shi et al. 2005). As mentioned in chapter 1, a second demethylase was identified shortly after, namely JmjC-containing histone demethylase 1 (JHDM1), which demethylates H3 mono- and diMeK36 through a reaction mechanism that is distinct from LSD1 (Yamane et al. 2006). More recently, a related protein named Jmjd2b has been shown to mediate the removal of tMeK9-H3 (Fodor et al. 2006), which could account for the tMeK9-H3 changes shown here.

The analysis of histone K20-H4 methylation also revealed unforeseen dynamics in the post-translational modification of histones that suggest an epigenetic role in cell differentiation. In particular, this involves an inverse relationship between mMeK20-H4 and tMeK20-H4, where the latter becomes progressively enriched in regions that correspond to differentiating cell populations – notably, those undergoing neurogenesis and myogenesis. Importantly, tMeK20-H4 levels have been shown to increase in aged tissue and growth arrested cells in comparison to logarithmically growing cells when assessed using chromatography and that this increase was not observed for mono- or di-methylated species (Sarg et al. 2002). Moreover, a recent study showed that tMeK20-H4 levels increase in serum-starved murine fibroblasts using an independently derived tMeK20-H4-specific antibody (Kourmouli et al. 2004). Taken together with the fact that histone K20-H4 methylation is not present in *S. cerevisiae*, but is found in higher eukaryotes such as *Drosophila*, mouse and human, this modification is likely involved in more complex developmental processes (Fang et al. 2002; Maison et al. 2002; Nishioka et al. 2002). In this regard, MeK20-H4 provides an important model system to evaluate the contribution of histone lysine methylation to the epigenetic control of cell differentiation and its influence on higher order chromatin organization during this process.

An examination of the distributions of K20-H4 methylation during neurogenesis has revealed important insight into the function of K20-H4-mediated chromatin dynamics and its role in development. The elevation of mMeK20-H4 in proliferating neural cells suggests this modification mediates a cell-cycle function that is not lineage restricted

because this distribution pattern was apparent throughout all other proliferating cell types in the embryo. The preferential enrichment of tMeK20-H4 in differentiated neurons found in the ventrolateral neuroepithelium indicates that this modification has a cell-type specific role that is associated with this neuronal pool, which is largely comprised of motor neurons. In support of a role for tMeK20-H4 in neuronal subtypes, a specific enrichment of this modification was observed in the differentiated chick ganglion cell layer at embryonic day 16. Importantly, as other retinal cell types are differentiated at this stage, this result demonstrates that tMeK20-H4 elevation is not a common feature of all neuronal differentiation events. We propose that tMeK20-H4 is involved in cell-type specific epigenetic regulation of differentiation and as such it will be important to identify the molecular players in this pathway.

Recent studies provide a connection between CRT and histone lysine methylation. CRT resides in the endoplasmic reticulum where it regulates the cytoplasm concentration of Ca^{2+}, which activates Calcineurin (CaN) at elevated levels of Ca^{2+}(Lynch et al. 2005). Activated CaN then phosphorylates myocyte enhancer factor 2C (MEF2C), a transcription factor important for specification of the primordial heart tube (Harvey 1999), inducing its translocation to the nucleus. In turn, MEF2C directly regulates the expression of Smyd1 (Phan et al. 2005), a SET domain gene specifically required for cardiac development (Gottlieb et al. 2002). Moreover, it appears that mice deficient for either CRT, MEF2C, or Smyd1 display similar cardiac pathology (Lynch et al. 2005; Phan et al. 2005), with a failure to undergo looping morphogenesis and right heart development (Lin et al. 1997). In this regard, it is possible that deregulated Smyd1 expression is responsible for the depletion

of tMeK20-H4 in CRT-deficient cardiomyocytes. Because the methyl-lysine specificity of Smyd1 if histone H3-K4 specific, it could potentially mediate an effect on tMeK20-H4 through indirect mechanism (Tan et al. 2006).

The majority of HKMTases are characterized by the presence of the SET/PR methyltransferase domain and there are potentially more than fifty of these enzymes encoded by the human genome (Arney et al. 2002). For the most part, however, it is not clear which lysine residue a given HKMTase modifies and this paucity also applies to how these enzymes catalyze the addition of mono-, di-, or tri-methyl moieties. For instance, ESET/SETDB1 has been shown to catalyze the formation of tMeK9-H3 from a di-methyl substrate (Fischle et al. 2003), yet it is not known which enzyme carries out the initial mono-methylation, and a similar scenario applies to *S. cerevisiae* Set1 (Santos-Rosa et al. 2002). Regardless of these gaps in mechanistic knowledge, there are clear biological consequences associated with HKMTase loss-of-function that support non-redundant roles for these enzymes during development (Yu et al. 1995; Peters et al. 2001; Tachibana et al. 2002; Erhardt et al. 2003; Rayasam et al. 2003). Importantly, such distinct functions are supported by the temporal and spatial differences in histone lysine methylation noted here, and by studies that analyze mono-, di-, and tri-methyl K9 histone H3 in cells lines lacking both *Suv39h1* and Suv39h2 or *G9a* HKMTases (Lehnertz et al. 2003). Finally, our results highlight the need to characterize patterns of histone lysine methylation *in vivo*, particularly in the emerging number of HKMTase knockout mice, which will greatly advance our understanding of HKMTase substrate specificities and the functional relationships between mono-, di-, or tri-methylation.

Chapter 4 ♦ Lineage specific genome-wide distributions and dynamics of histone lysine methylations

Portions of this chapter have been published or are being prepared for publication in:

1. **Vincent L. Biron**, Kirk J.McManus, Ninghe Hu, Michael J. Hendzel and D. Alan Underhill. (2004) Distinct dynamics and distribution of histone methyl-lysine derivatives in mouse development. *Dev Biol* **276 (2)**, 336-50

2. **Vincent L. Biron** and D. Alan Underhill. Genome-wide distributions and dynamics of histone methyl-lysine derivatives in skeletal muscle differentiation. *In preparation*

Introduction

Lysine methylation of histones is associated with distinct chromatin domains and is dependent on the position and state of the modified residue. For instance, tMeK9-H3, mMeK27-H3 and tMeK20-H4 are associated with pericentric heterochromatin, tMeK27-H3 and mMeK20-H4 are associated with facultative heterochromatin, and tMeK4-H3 is found exclusively in active euchromatin. These modifications have been shown to dynamically regulate gene promoters during development and also display cell-type variability with respect to their association with repetitive elements (Margueron et al. 2005; Martens et al. 2005). The role of these modifications *in vivo* is not entirely understood and requires a comprehensive genome-wide description by imaging and chromatin immunoprecipitation approaches. In this chapter, the *in situ* distributions of histone methyl-lysine residues were investigated during cell differentiation using myogenic models.

The concept that chromatin modifying-enzymes are essential modulators of differentiation stems largely from studies using skeletal muscle cell lines. These cell lines can be grown under conditions of proliferation and induced to undergo robust differentiation. Importantly, the molecular mechanisms that have been uncovered using skeletal muscle cultures are applicable to *in vivo* systems (McKinsey et al. 2001; McKinsey et al. 2002). In this regard, much of what is known about chromatin dynamics in myogenesis pertains to changes in histone acetylation. For instance, the myogenic regulators Myod and MEF2 have been shown to interact with HATs and HDACs for the activation and repression of target genes respectively (McKinsey et al. 2001). By directing

the chromatin modifying activities of HATs and HDACs, Myod and MEF2 can differentially regulate gene expression in a context dependent manner to regulate gene programs that dictate either myocyte proliferation or differentiation. In addition, the expression of class II HDACs 4, 5, 7 and 9 is enriched in the heart, skeletal muscle, and brain, while class I and III HDACs appear to be ubiquitous (McKinsey et al. 2002). This would suggest that these cell lineages may have unique genome deacetylation requirements, which may extend to other histone modifications. Consistent with this idea, these three lineages are characterized by visible changes in histone H4-K20 methyl derivatives during mouse embryogenesis (Chapter 3).

A role for histone lysine methylation in myogenesis is emerging. In early muscle differentiation, the myogenin promoter has been shown to be repressed through histone H3-K9 methylation and is later activated by acK9-H3 (McKinsey et al. 2002). MEF2 regulates this promoter by coordinated regulation of acetylation and methylation, a process that occurs through HP1-HDAC/ HP1-Suv39H1 HKMTase interactions (Yamane et al. 2006). The histone methyltransferase Ezh2 has also been shown to repress the myogenic program through methylation of histone H3-K27, with loss upon differentiation (Caretti et al. 2004). Therefore, there appears to be important regulatory events in myogenesis that are dependent on histone lysine methylation, which function in the epigenetic maintenance of this lineage.

Recently, we have shown that in the myotome and neuroepithelium of the mouse embryo, the levels of mMeK20-H4 are progressively lost with a concomitant increase in tMeK20-H4 over the course of differentiation (Biron et al. 2004). In this chapter, we have

examined a broad range of histone methyl-lysine derivatives during myogenesis. Using mouse C2C12 cells and primary limb bud cultures, 14 histone methyl-lysine derivatives are examined in skeletal myogenesis. In addition to further refining changes in K20-H4 derivatives, we found that methyl modifications involving K20-H4, K36-H3 and K79-H3 exhibited differences as cells progressed from myoblasts to multinucleated myotubes.

Results

Histone lysine methylation dynamics during skeletal muscle differentiation

We examined the distribution of 14 methyl-lysine antibodies in C2C12 cells following 1, 3, or 7 days of serum withdrawal-induced differentiation. C2C12 cells are mouse skeletal muscle cells, which have have been widely used to study muscle differentiation. Gross microscopic examination of interphase nuclei over the course of differentiation revealed methyl derivatives of histone H3 and H4 underwent nuclear-wide changes, while other methyl-lysine species appeared relatively static. In each case, three-dimensional immunolocalization data was obtained with respect to DAPI-rich heterochromatin domains, over the course of differentiation. Importantly, this well established marker also provides a standard for nuclear staining, taking into account the amount of heterochromatin of each nucleus and the cell cycle stage.

Static methyl-lysine derivatives

The interphase distributions mMeK4-H3, dMeK4-H3, tMeK4-H3, dMeK23-H3, mMeK9-H3, tMeK9-H3 and tMeK9-H3S10p antibodies displayed relatively static distributions over the course of differentiation. All of these antibodies with the exception of tMeK9-H3 and

tMeK9-H3S10P were enriched in domains outside of DAPI-rich pericentric heterochromatin. A detailed analysis of the distributions of tMeK4-H3, dMeK23-H3, tMeK9-H3 and K9-H3S10p antibodies is shown in Figure 4-1. Based on these immunofluorescence findings, we suggest these methyl-lysine derivatives are not involved in significant genome-wide chromatin rearrangements during muscle differentiation. It is possible, however, that these modifications are involved in locally dynamic processes that are not detectable in our analysis.

Dynamic distributions of H4 K20 methylation

The three dimensional distributions of histone H4 mono-, di- and tri-methyl K20 displayed significant dynamics over the course of skeletal muscle differentiation (Figure 4-2). Importantly, these antibodies initially detected chromatin structures consistent with previous reports that used different antibodies directed against the same antigens (Kourmouli et al. 2004; Schotta et al. 2004). In undifferentiated C2C12 myoblasts at day 1, mMeK20-H4 was found in two large non-heterochromatic domains that coincided with the nuclear and nucleolar periphery in > 95% of interphase cells, as well as in hundreds of smaller foci found throughout the nucleoplasm (Figure 4-2 A). However, during the course of differentiation (day 3-7), there was a dramatic loss of the large mMeK20-H4 accumulations, while smaller foci were retained. The disappearance of large mMeK20-H4 accumulations also coincided with an overall depletion of mMeK20-H4 levels as detected by immunoblot analysis (Figure 4-5). As a result, the highest levels of mMeK20-H4 are associated with the proliferative state at day 1, which is consistent with an elevation of H3S10p, a modification associated with mitotic chromosomes (Figure 4-5). Although

dMeK20-H4 levels were relatively static by immunoblot analysis (Figure 4-5), its subnuclear distribution was altered during C2C12 differentiation (Figure 4-2 B). In ~ 80 % of proliferating myoblasts at day 1, an enrichment of dMeK20-H4 was found in nucleoli (Figure 4-2 B, arrows), and this distribution was completely lost by day 3 of serum withdrawal. Interestingly, all nuclei with dMeK20-H4 nucleolar enrichment were positive for Myod, a central regulator of myogenesis, but not all Myod-positive nuclei had nucleolar dMeK20-H4 (Figure 4-2 D). This could suggest a link between MyoD levels and chromatin changes that result in nucleolar enrichment of dMeK20-H4. As noted previously, tMeK20-H4 remained enriched in pericentric heterochromatin domains as defined by DAPI, which increased in size over the course of differentiation (Figure 4-2 C). Consistent with this observation, tMeK20-H4 levels increased during differentiation when analyzed by immunoblot analysis. Significantly, this analysis reveals gross changes in histone K20-H4 methylation status during skeletal muscle differentiation and suggests distinct roles for the mono-, di- and tri-methyl modifications in the epigenetic regulation of muscle differentiation.

Dynamic distributions of H3 K36 methylation

The subnuclear distributions of mMeK36-H3 and tMeK36-H3 remained relatively static over the course of differentiation, maintaining punctate distributions excluded from DAPI-rich heterochromatin (Figure 4-3), and their levels did not change appreciably during differentiation when monitored by immunoblot analysis (Figure 4-5 A). However, in differentiated cells at days 3 and 7, some nuclei had an absence of mMeK36-H3. In addition, nuclei within the same myotube had dramatic differences in their levels of

mMeK36-H3. Given the static levels of mMeK36-H3 in immunoblot blots, the overall change in global nuclear mMeK36-H3 immunodetection may represent a chromatin reorganization event generating a masked epitope during differentiation. However, the number of nuclei displaying an absence of mMeK36-H3 may not be within the detection range of immunoblot analysis – 1.9% depletion at day3 and 17.8% depletion at day 7. Nevertheless, the strikingly different levels of mMeK36-H3 within nuclei of the same myotube would indicate they might have different epigenetic profiles.

Dynamic distributions of H3 K79 methylation

Immunodetection of mMeK79-H3 and tMeK79-H3 was dramatically reduced in the nucleus of differentiated C2C12 cells, but revealed an elevated signal in the cytoplasm (Figure 4-4). In contrast, dMeK79-H3 levels retained a relatively static genome-wide distribution with no cytoplasmic signal over the course of differentiation (Figure 4-4 B). Nevertheless, the levels of mMeK79-H3 and tMeK79-H3 appeared to remain static when measured by immunoblot analysis or nuclear protein (Figure 4-5 A), indicating that antigens are preserved. The disparity in detection of these modifications by immunofluorescence and immunoblot blotting may be explained by an epitope masking event occurring either through a reorganization of m/dMeK79-H3-rich chromatin, or tenacious protein binding to these domains arising during skeletal muscle differentiation.

The cytoplasmic localization of mMeK79 and tMeK79 antibodies displayed a repetitive banding pattern reminiscent of the distribution of myosin heavy chain (MHC) in muscle sarcomeres. We therefore examined the relative localization of MHC, with respect to mMeK79-H3 and tMeK79-H3. Extensive cytoplasmic co-localization was observed

between MHC and mMeK79-H3 and tMeK79-H3 in multinucleated C2C12 cells (Figure 4-4). Immunoblot blots comparing cytoplasmic C2C12 extracts from proliferating versus differentiating cells indicate prominent reactivity of mMeK79-H3 and tMeK79-H3 with a non-histone protein in differentiated extracts that could correspond with an MHC-specific band (Figure 4-4 B). In contrast, tMeK4-H3, mMeK20-H4, and tMeK20-H4 antibodies did not detect any protein in C2C12 whole cell extracts, which do not contain concentrated histones detected by these antibodies (data not shown). These results nevertheless suggest that mMeK79-H3 and tMeK79-H3 antibodies detect a similar lysine methylation epitope associated with the sarcomere that is elevated upon muscle differentiation. This could possibly represent the signature of a histone methyltransferase with dual specificity for nuclear and cytoplasmic lysine methylation targets. One such HKMTase may be Ezh2, which methylates K27-H3 and also associates with cytoplasmic actin (Yamane et al. 2006), a protein that is also enriched in sarcomeres.

K20-H4 methylation analysis in mouse primary cultures

We next examined whether the K20-H4 methylation dynamics seen in skeletal myogenesis is a common property of all differentiating cells types (see Chapter 3). We therefore investigated the dynamics of histone H4-K20 methylation in a mouse primary culture system that contains a range of differentiating cell types. During the initial phase of these cultures, mitotically active myoblasts are present and eventually differentiate into multinucleated myotubes, which can be discerned by light microscopy as elongated cells with 3 or more nuclei by approximately day 3. In addition, chondrogenic precursors aggregate and form visible nodules as early as day 2. This system therefore enabled a more

rigorous investigation of K20-H4 methylation levels over the course of myogenesis in comparison to non-muscle lineages in the same culture. For mMeK20-H4, although its levels were highly variable between cells, its signal intensity in multinucleated myotubes was consistently at the lower end of this range at both day 5 and day 7 of culture (Figure 4-6 A, B). In contrast, tMeK20-H4 levels were noticeably elevated in day 5 myotubes and increased further by day 7 (Figure 4-6 C, D). As a control, mMeK4-H3, tMeK4-H3 and mMeK9-H3 levels were not enriched in any particular cell lineage, while tMeK9-H3 was markedly elevated in mitotic chromosomes, consistent with their behavior in embryos (Chapter 3). Moreover, there was no connection between histone K20-H4 methylation and cell differentiation in chrondrocytic nodules where both the mono- and trimethylated states were randomly elevated from day 3-7 (Figure 4-7), further supporting a specific role for K20-H4 methylation in myogenesis. Slight fluctuations in chrondrocytic nodules may be due to cell cycle related variations that occur throughout the primary culture (Figure 4-7). As expected, mMeK20-H4 levels are more often elevated in bromodeoxyuridine (BrdU) positive nuclei, whereas tMeK20-H4 levels are lower (Figure A-2). The relationship between K20-H4 methylation and myogenesis was corroborated by examining mMeK20-H4 and tMeK20-H4 signal intensities in MF20-positive muscle cells (Figure 4-6 E, F). In 59 randomly scored MF20-positive cells (day 3), all nuclei had mMeK20-H4 levels in the lowest 33% of signal intensity (Figure 4-6 G), which was in marked contrast to tMeK20-H4 where levels were in the upper two-thirds of intensity values in 98% (44/45) of MHC-positive cells. These data indicate that, while global mMeK20-H4 levels decrease in differentiating myotubes, tMeK20-H4 levels increase, and therefore parallels their behavior

in developing mouse neuroepithelium and striated muscle (Chapter 3). Taken together, the unique behavior of histone H4-K20 methyl derivatives during myogenesis suggests that these modifications play a lineage specific role in differentiation.

Figure 4- 1: Immunfluorescence of select histone lysine methylations showing static distributions throughout C2C12 muscle differentiation.

Immunofluorescence in C2C12 cells following 1, 3, or 7 days of differentiation induced by serum withdrawal. Antibody stain is shown in left panels, DAPI stain in center panels and merged in right panels (antibody=red, DAPI=green). The three dimensional distributions of methylations relative to antibodies are shown in the z-y and z-x axes on the bottom and to the right of each image panel, respectively. A) tMeK9-H3, B) tMeK9-H3S10p, C) tMeK4-H3 and D) dMeK23-H3, all showing genome wide distributions not associated with pericentric heterochromatin. Note the distribution of tMeK9-H3 with this antibody is not consistent with localization reported in the literature as shown in Figure 3-2, but more closely resembles the distribution of dMeK9-H3.

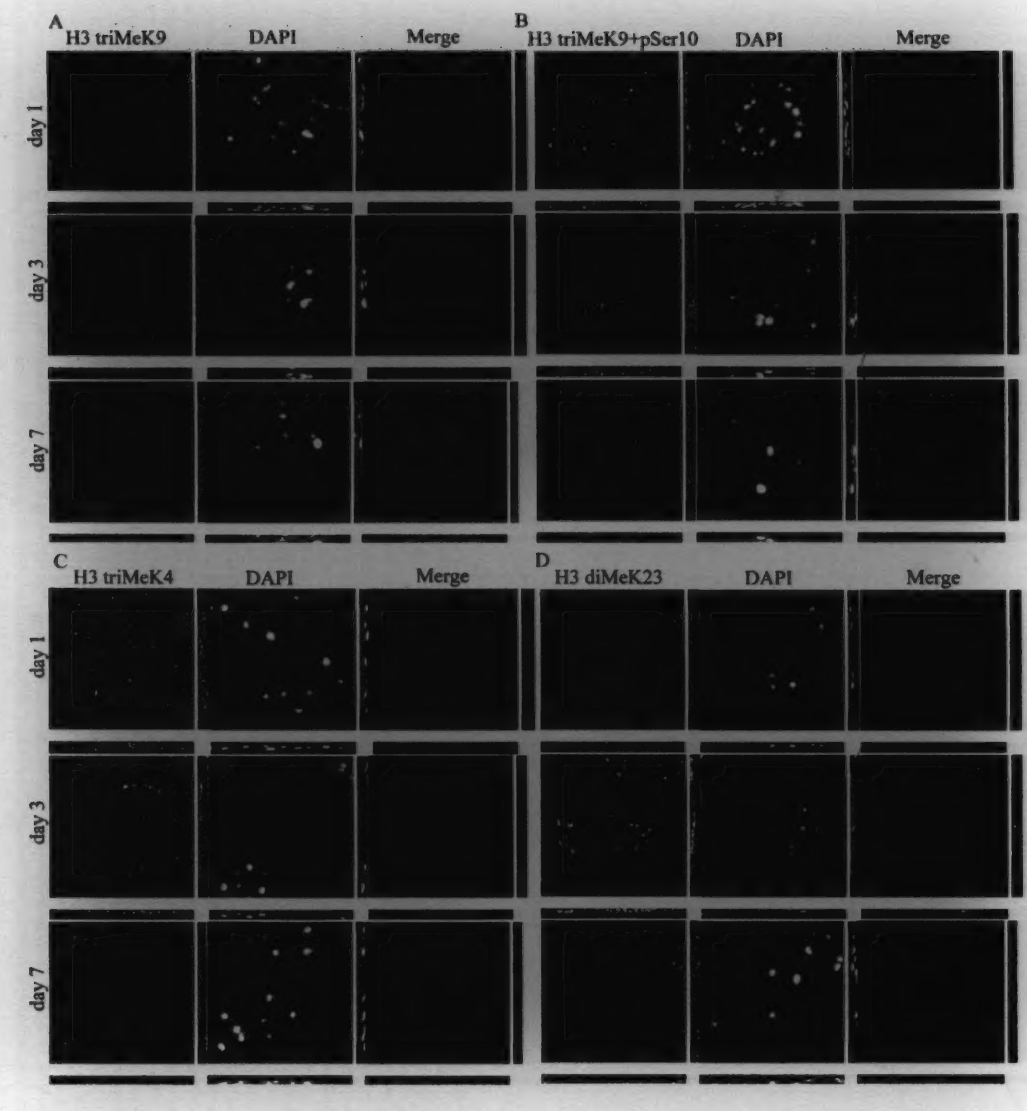

A
H3 triMeK9
DAPI
Merge
B
H3 triMeK9+pSer10
DAPI
Merge
day 1
day 3
day 7
C
H3 triMeK4
DAPI
Merge
D
H3 diMeK23
DAPI
Merge
day 1
day 3
day 7

Figure 4- 2: Immunofluorescence analysis of MeK20-H4 over the course of skeletal muscle differentiation.

Immunofluorescence in C2C12 cells following 1, 3, or 7 days of differentiation induced by serum withdrawal. Antibody stain is shown in left panels, DAPI stain in center panels and merged in right panels (antibody=red, DAPI=green). The three dimensional distributions of methylations relative to antibodies are shown in the z-y and z-x axes on the bottom and to the right of each image panel, respectively. A) mMeK20-H4 levels progressively decrease over the course of differentiation, with large aggregates (arrows) disappearing during differentiation. C) In contrast, tMeK20-H4 levels progressively increase. B) An enrichment of dMeK20-H4 is seen in nucleoli of proliferating cells (day1), which is not seen in differentiating cells. D) Immunolocalization of dMeK20-H4 (red), with respect to Myod (green) and DAPI (blue), merged in the rightmost panel.

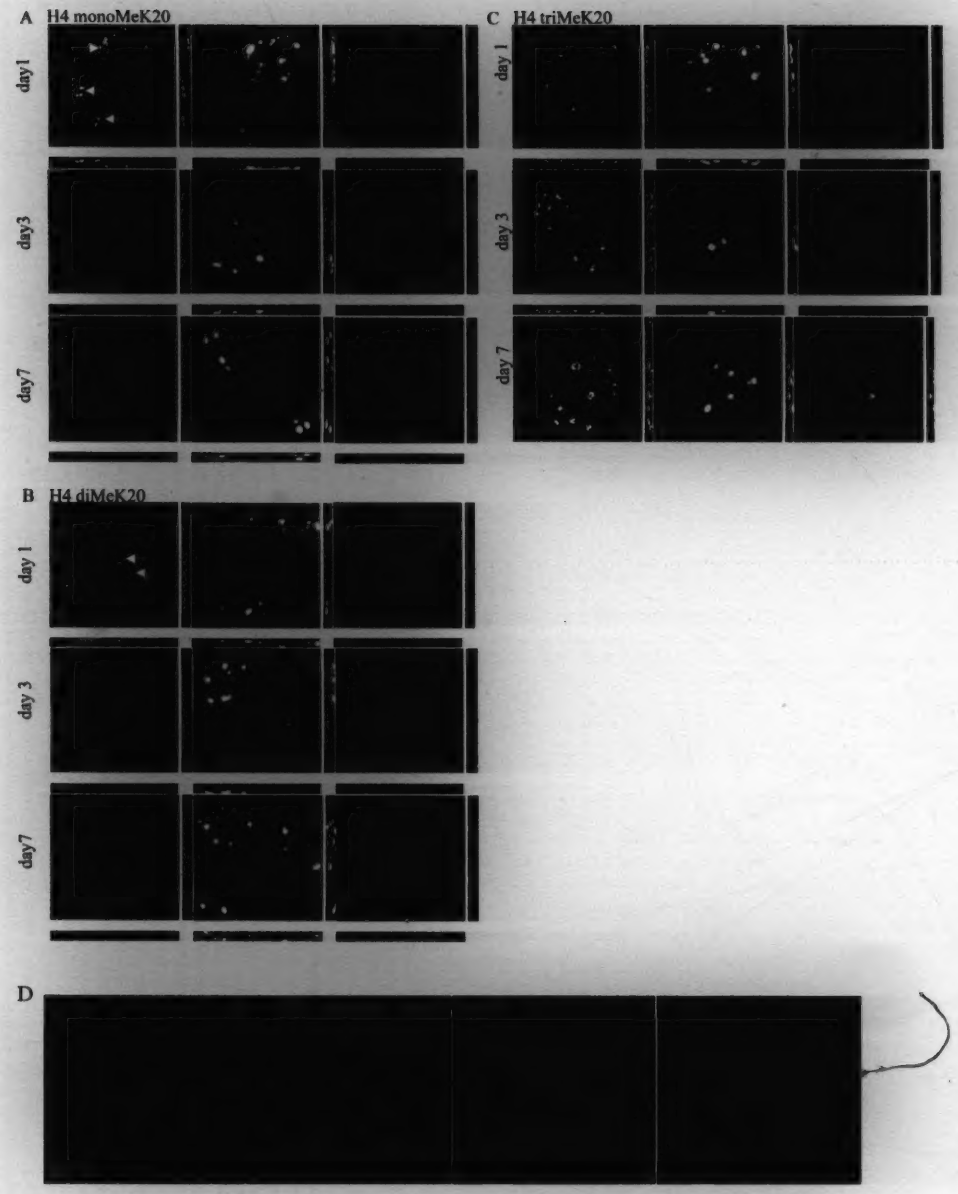

A H4 monoMeK20
day1
day3
day7
C H4 triMeK20
day 1
day 3
day 7
B H4 diMeK20
day 1
day 3
day7
D

Figure 4- 3: Immunofluorescence analysis of MeK36-H3 over the course of skeletal muscle differentiation.

Immunofluorescence in C2C12 cells following 1, 3 or 7 days of differentiation induced by serum withdrawal. Antibody stain is shown in left panels, DAPI stain in center panels and merged in right panels (antibody=red, DAPI=green). The three dimensional distributions of methylations relative to antibodies are shown in the z-y and z-x axes on the bottom and to the right of each image panel, respectively. A) Subnuclear distribution of mMeK36-H3 is maintained with respect to DAPI, however some nuclei are depleted of this methylation B) following 7 days of differentiation, with significant differences in levels between nuclei of the same myotube. C) tMeK36-H3 remains relatively unchanged over the course of differentiation. D) Quantification of mMeK36-H3 depletion in C2C12 nuclei. The percentage of nuclei depleted for mMeK36-H3 in mononuclear and polynuclear cells was obtained from 265 nuclei at day 3 and 163 nuclei at day 7 of differentiation.

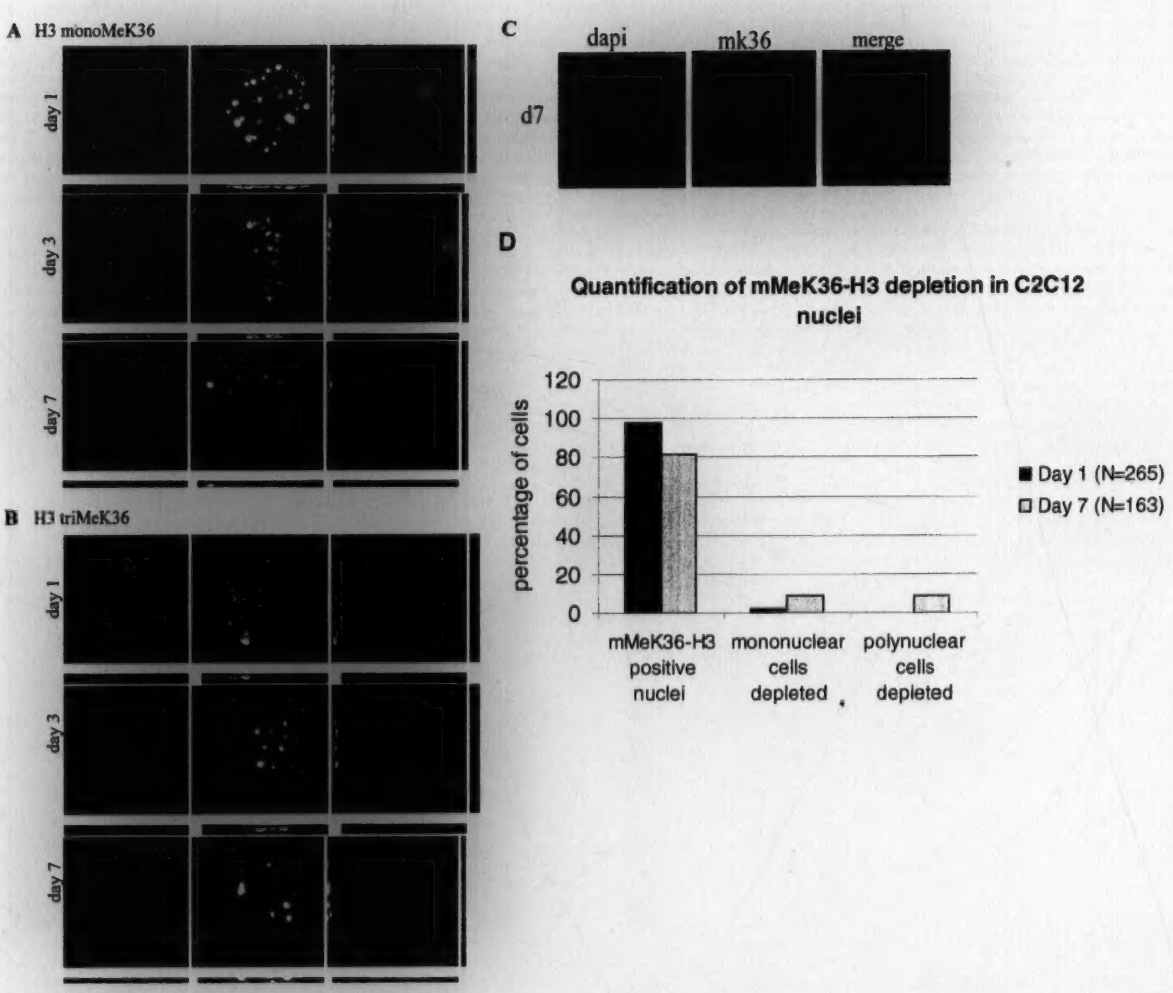

A H3 monoMeK36
day 1
day 3
day 7
B H3 triMeK36
day 1
day 3
day 7
C
dapi
mk36
merge
d7
D
Quantification of mMeK36-H3 depletion in C2C12 nuclei
percentage of cells
120
100
80
60
40
20
0
mMeK36-H3 positive nuclei
mononuclear cells depleted
polynuclear cells depleted
Day 1 (N=265)
Day 7 (N=163)

Figure 4- 4: Immunofluorescence analysis of MeK79-H3 over the course of skeletal muscle differentiation.

Immunofluorescence in C2C12 cells following 1, 3, or 7 days of differentiation induced by serum withdrawal. Antibody stain is shown in left panels, DAPI stain in center panels and merged in right panels (antibody=red, DAPI=green). The three dimensional distributions of methylations relative to antibodies are shown in the z-y and z-x axes on the bottom and to the right of each image panel, respectively. At day 1, m/d/tMeK79-H3 are distributed in the nucleoplasmic interior excluded from heterochromatin (A-C). mMeK79-H3 and tMeK79-H3 levels are significantly depleted upon differentiation (B, C) and antibodies co-localize with myosin heavy chain (MF20).

A
H3 monoMeK79
DAPI
Merge
day 1
day 3
B
H3 diMeK79
DAPI
Merge
day 1
day 3
day 7
C
H3 triMeK79
DAPI
Merge
day 1
day 3

Figure 4- 5: Immunoblot analysis of select histone methyl-lysines in differentiating C2C12 cells.

Immunoblot blot analysis of acid extracted histones from C2C12 cells over the course of differentiation. A coomassie stain is shown to indicate histone levels. A progressive increase in tMeK20-H4 and decrease in mMeK20-H4 is seen over the course of differentiation. Relatively static levels of mMeK79-H3, tMeK79-H3 and mMeK36-H3 in immunoblot blots may indicate chromatin changes *in vivo* that alter antibody epitope accessibility in immunofluorescence. H3S10p is detected as an indicator of proliferation.

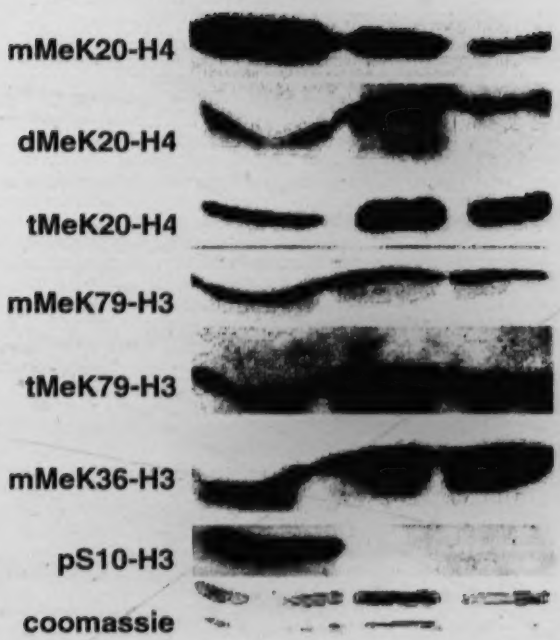
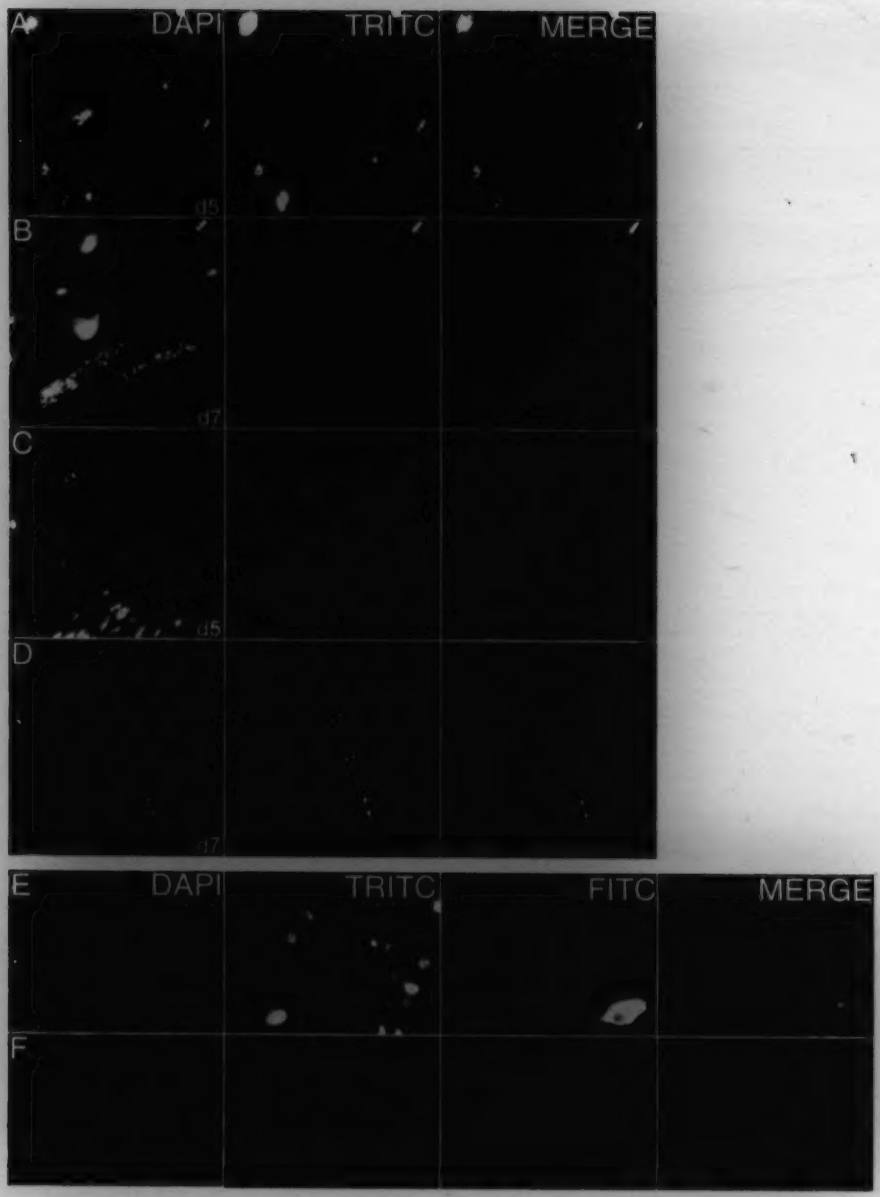

mMeK20-H4
dMeK20-H4
tMeK20-H4
mMeK79-H3
tMeK79-H3
mMeK36-H3
pS10-H3
coomassie

Figure 4- 6: Muscle specific changes in H4 mono- and trimethyl lysine 20 levels in mouse primary limb bud cultures.

Mouse primary limb bud cultures derived from E10.5 embryos were analyzed at day 5 (A, C) and day 7 (B, D) for the distribution of mMeK20-H4 (A, B) and tMeK20-H4 (C, D) (TRITC column) with respect to DAPI (left column). The relationship between (E) mMeK20-H4 and (F) tMeK20-H4 (TRITC), DAPI and MF20 (FITC) positive cells is shown in merged images (red: TRITC, blue: DAPI, green: FITC). (G) MF20 positive nuclei were analyzed in twenty-five 40x field images (as shown in E and F) for the relative intensity distributions of mMeK20-H4 and tMeK20-H4, which were scored as being in either the lower, middle, or upper third of intensity values.


A
DAPI
TRITC
MERGE
d5
B
d7
C
d5
D
d7
E
DAPI
TRITC
FITC
MERGE
F

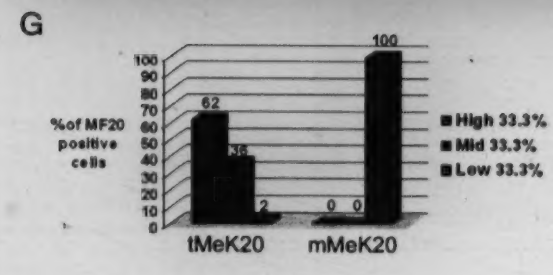
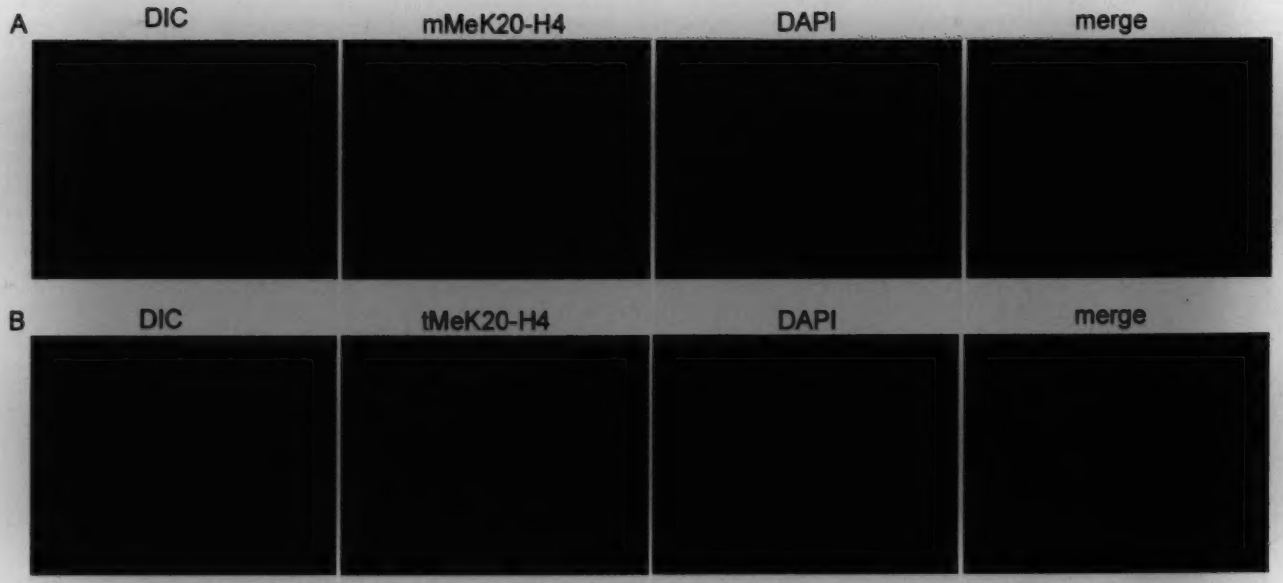

G
%of MF20 positive cells
100
90
80
70
60
50
40
30
20
10
0
62
35
2
0
0
100
tMeK20
mMeK20
High 33.3%
Mid 33.3%
Low 33.3%

Figure 4- 7: Histone H4-K20 methylation in differentiating chondrocytes derived from mouse primary limb bud cultures.

Differentiating chrondocytes in primary cultures from E10.5 mouse limb buds are shown in nodules of DIC images. Cultures were stained by immunofluorescence with A) mMeK20-H4 or B) tMeK20-H4 with respect to DAPI. Merged images are shown in rightmost panels (antibody=red, DAPI=green). Trimethylated H4-K20 appears to be preferentially elevated at the periphery and mMeK20-H4 in the center of the nodule, however, this is due to the plane in which these images are taken in an elevated nodule.


A
DIC
mMeK20-H4
DAPI
merge
B
DIC
tMeK20-H4
DAPI
merge

Discussion

Distinct histone methyl-lysine residues are dynamic in skeletal muscle differentiation

Several studies have examined global gene expression patterns over the course of C2C12 myogenesis (Moran et al. 2002; Shen et al. 2003; Tomczak et al. 2004). Although different experimental approaches were taken, these studies all revealed that the transition from myoblast proliferation to differentiation is characterized by large-scale gene activation and repression events. This study provides a nuclear imaging perspective to the global gene regulation dynamics occurring during C2C12 myogenesis.

Euchromatic rearrangements during differentiation

Methylations of lysines 4, 36 and 79 on histone H3 are hallmarks of active chromatin and have non-overlapping genic distributions that are thought to coordinate distinct transcriptional events (Farris et al. 2005). Specifically, d/tMeK4-H3 is enriched at the 5' end of genes and promotes initiation (Schneider et al. 2004), a gradient of MeK79-H3 that decreases 5 'to 3' is associated with elongation, and d/tMeK36-H3 elevation at the 3' end is thought to regulate elongation-termination and possibly co-transcriptional splicing (Krogan et al. 2003; Bannister et al. 2005; Kizer et al. 2005). Consistent with their roles in transcriptional activation, the subnuclear distributions of tMeK4-H3, d/tMeK79-H3 and tMeK36-H3 are not associated with DAPI-rich heterochromatin (Figures 4-1 C, 4-3 and 4-4). While the transcriptional function of monomethylated species of these residues is not clear, their distributions shown here suggest that these modifications are largely excluded from constitutive heterochromatin domains.

The marked difference in immunofluorescence intensity of mMeK36-H3 in nuclei of the same myotube in differentiated muscle could have interesting implications. This disparity between nuclei could suggest there are epigenetic differences between nuclei exposed to the same cytoplasmic environment. Heterokaryon analysis suggests that a critical dosage of nuclei in muscle cells may be required such that an elevation of trans-acting factors could negatively affect the rate of gene expression (Pavlath and Blau 1986). The finding that mMeK36-H3 levels are depleted in a subset of terminally differentiated myotubes (Figure 4-3) may be a consequence of muscle cell fusion whereby cytoplasmic factors initially repress mMeK36-H3 methylation until the cell adjusts its levels appropriate to the number of nuclei present. Alternatively, these depletions may represent a cell-cycle mediated event of chromatin condensation/decondensation. At present, the function of mMeK36-H3 is speculative but our analysis suggests it has a unique epigenetic role during myogenesis.

We also observed depletion in nuclear staining of mMeK79-H3 and mMeK79-H3 upon differentiation. The decreased nuclear signal could be the result of a protein binding event that masks the mMeK79 and tMeK79-H3 epitopes, as our immunoblot blot analysis suggests these antigens remain relatively constant throughout differentiation. In yeast, Sir proteins bind to K79-H3 in telomeric silencing to prevent its methylation (Yamane et al. 2006). In the DNA damage response, 53BP1 is recruited to DSBs through dMeK79-H3 via its Tudor domains (Huyen et al. 2004). As the levels of dMeK79-H3 remain unchanged following DNA damage, 53BP1 recruitment occurs by increasing the accessibility of

dMeK79-H3, an event that is thought to be attributed to a loss in higher order chromatin structure (Huyen et al. 2004). In this regard, it is possible that an apparent loss in mMeK79 and tMeK79-H3 is a result of higher order chromatin folding events that mask these epitopes. Consistent with such an event are our observations of accumulating heterochromatin-associated histone H4 methyl-lysines and changes in DAPI staining over the course of skeletal muscle differentiation.

The increased sarcomere-associated signal of mMeK79-H3 and tMeK79-H3 antibodies during muscle differentiation could represent the signature of an HKMTase with both nuclear and cytoplasmic targets. The only enzyme with known specificity for MeK79-H3 is Dot1 and it is not currently known whether this HKMTase has cytoplasmic targets. Nevertheless, it is becoming increasingly apparent that HKMTases, like PRMTases, can methylate non-histone targets. For instance, Set7/9 can specifically lysine methylate p53 and TAF10 transcription factors (Chuikov et al. 2004; Kouskouti et al. 2004). Ezh2 can methylate histone H3-K27 in the nucleus, and can also regulate actin polymerization in the cytoplasm where it is dependent on functional HKMTase activity (Caretti et al. 2004; Yamane et al. 2006). Notably, α-actin levels are dramatically upregulated in C2C12 cells following 2 days of differentiation (Shen et al. 2003) and its sequence contains similarity to the histone H3-K79 flanking region which could present a similar epitope if it were lysine methylated. Interestingly, the Ezh2 HKMTase methylates tMeK27-H3 during muscle proliferation, and upon differentiation dissociates from its targets (Caretti et al. 2004). It is tempting to speculate that this dissociated complex may function in cytoplasmic actin polymerization during muscle differentiation. Cross-talk between the sarcomere and gene

regulation have been suggested, but the underlying mechanisms remain elusive (Kuwahara et al. 2005). Taken together, it is possible that the sarcomere-associated elevation in mMeK79 and tMeK79 antibody signals represent sarcomere protein methylation by the action of a protein lysine methyltransferase that remains to be identified. Like protein arginine methylation, lysine methylation could be a more widespread mechanism of protein regulation whereby its methyltransferases may have both histone and non-histone targets.

Dynamic K20-H4 methylation during differentiation

Our analysis of histone H4-K20 methylation demonstrates that all three methylation states on this residue undergo marked changes during skeletal muscle differentiation. Upon C2C12 differentiation, we observed a loss of monomethyl-rich nuclear aggregates and dimethylated nucleolar enrichments, with a progressive increase in pericentromeric heterochromatin-associated trimethylation. The concomitant dynamics of all three methylation states suggest interdependence between some of these modifications, consistent with previous reports (Julien and Herr 2004; Karachentsev et al. 2005) and results in Chapter 3. Moreover, our results suggest that each methylation state of histone H4-K20 confers distinct chromatin states. Given the proposed ability of the H4 tail to participate in inter-nucleosomal interactions (Luger et al. 1997), it is possible that each MeK20-H4 derivative provides unique structural properties that enable distinct higher-order chromatin folding or protein binding events.

The increase in tMeK20-H4 observed in myotubes may facilitate the formation and/or maintenance of pericentromeric heterochromatin during differentiation. This is consistent with previous immunofluorescence and electron microcopy analyses of rat L6E9

muscle cells where centromeric heterochromatin underwent an increase in size and relocated to more peripheral regions of the nucleus in the myoblast (day 1) to myotube (day 7) transition (Chaly and Munro 1996). Likewise, previous biochemical experiments have shown that there are gross structural changes in chromatin during chick muscle differentiation (Wiid et al. 1988), and in the differentiation of various neural lineages (Kuenzle et al. 1983). Recently, pericentromeric heterochromatin domains have also been observed to undergo large-scale movements that are specific the Purkinje cells (Solovei I et al 2004). Our results are therefore consistent with a model whereby tMeK20-H4 contributes to changes in chromatin structure that are required for – or contribute to – cell differentiation during development. In this capacity, it is interesting to note that components of the tMeK9-H3 pathway are required to mediate silencing of cell cycle control genes during C2C12 differentiation (Ait-Si-Ali et al. 2004). Taken with the suggestion that the tMeK9-H3 and tMeK20-H4 marks may be interdependent (Kourmouli et al. 2004), the localization of tMeK20-H4 to pericentromeric heterochromatin may reflect a role in the stable silencing of genes that are not involved in the maintenance or function of differentiated cell-types.

Multiple lines of evidence suggest an important role for K20-H4 monomethylation in cell cycle regulation. Recently, Pr-Set7 was shown have high specificity for the catalysis of mMeK20-H4 (Xiao et al. 2005). Independent studies have shown that Pr-Set7 regulates K20-H4 methylation levels throughout the cell cycle, which is required for mitosis (Rice et al. 2002; Karachentsev et al. 2005). Consistent with these findings, we show that proliferating myoblasts have elevated mMeK20-H4 that is significantly depleted upon

mitotic exit and cell differentiation. In addition, the distribution of mMeK20-H4 in proliferating cells is enriched in two large domains with irregular morphology resembling chromosome territories, an association that is lost as cells exit the cell cycle and undergo differentiation. Importantly, these accumulations have been reported to represent the inactive X (X_i) chromosome (Kohlmaier et al. 2004; Schotta et al. 2004). In support of the X_i having a function in myoblast proliferation, Ezh2-dependent trimethylation of K27-H3, also a mark of X_i, is lost during myoblast differentiation (Caretti et al. 2004). Moreover, Pr-Set7 suppressor of variegation activity has been shown to be associated with *Drosophila* chromosome 4, an autosome that has characteristics of the X_i (Larsson et al. 2001; Karachentsev et al. 2005). Interestingly, the persistence of histone H4-K20 monomethylation in Pr-Set7 deficient *Drosophila* cells suggests this modification can be stable over several days and cell generations (Karachentsev et al. 2005). The loss of monomethylation upon differentiation observed in culture could therefore be due to either conversion of this modification to higher methylation states, the action of a demethylase, which remains to be identified for histone H4-K20, or nucleosome replacement. Taken together, mMeK20-H4 appears to play an epigenetic and cell-cycle regulatory role that is chromosome specific. This function may be manifest through heterochromatin silencing of cell-cycle genes, checkpoint control or regulation of chromosome segregation, as suggested by Karachentsev et al. (Karachentsev et al. 2005).

The nucleolar enrichment of dMeK20-H4 in Myod positive myoblasts provides additional clues to histone H4-K20-mediated gene regulation during differentiation. As the levels of dMeK20-H4 were not measurably altered during differentiation, this transient

enrichment may underlie chromatin rearrangements involving loci coordinately regulated by the nucleolus. For example, ribosomal DNA loci are localized to the nucleolus where they are spatially regulated. A transient enrichment of dMeK20-H4 in this region could repress genes involved in muscle differentiation as the cells are proliferating. It is worth noting here that we have also observed an enrichment of histone H3.3 in the nucleolus (data not shown), with a distribution pattern similar to that of dMeK20-H4, however, the localization of this H3.3 antibody does not correspond with localization of H3.3-GFP (M.J. Hendzel, personal communication).

The progressive elevation of tMeK20-H4 associated with pericentromeric heterochromatin during muscle differentiation is consistent with several reports and has important implications. During skeletal muscle differentiation, pericentromeric heterochromatin foci coalesce in a DNA-methylation-dependent process, thereby increasing in size through large scale chromatin rearrangements (Brero et al. 2005). In Purkinje cell post-natal differentiation, large scale movements of pericentromeric heterochromatin have also been described (Solovei et al. 2004) Interestingly, although both tMeK9-H3 and tMeK20-H4 have been shown to have an association with pericentromeric heterochromatin, we have shown that only tMeK20-H4 levels increased over the course of skeletal muscle differentiation (Shown with 2 distinct tMeK9-H3 antibodies, Chapters 3/4). Although the enzymes that catalyse histone H4-K20 and H3-K9 trimethylation are linked (Schotta et al. 2004), distinct enrichments of tMeK9-H3 and tMeK20-H4 can occur in certain repetitive elements (e.g. IAP LTR and Charlie) (Martens et al. 2005). This would suggest that tMeK9-H3 and tMeK20-H4 can behave autonomously as chromatin repressive

marks. The apparent coexistence between these two marks by immunofluorescence may be limited by resolution, such that both marks are found in heterochromatin of distinct subtypes. Our data demonstrating tMeK20-H4 changes that are independent of tMeK9-H3 levels suggest that these modifications can mediate independent epigenetic roles in muscle development. It would therefore be interesting to trace repeat-associated lysine methylation changes over the course of skeletal muscle differentiation.

Various models of cell differentiation have been proposed to explain its apparent unidirectional nature, with contrasting views of progressive gene repression versus gene activation (Fisher and Merkenschlager 2002). Microarray data following the expression of thousands of genes during C2C12 differentiation demonstrate that large-scale changes in expression occur, which include both gene repression and activation events (Shen et al. 2003; Tomczak et al. 2004). Our data are consistent with these studies, supporting a model whereby myogenesis is characterized by progressive genome-wide changes in repression and activation, depending on the chromatin regions examined. For instance, mMeK20-H4 rich genes may be repressed during proliferation, but are activated upon differentiation, whereas some active genes in cycling myoblasts could become repressed by tMeK20-H4 during differentiation. Importantly, various lines of evidence would suggest that histone H4-K20 repressive marks are highly stable (Karachentsev et al. 2005; Martens et al. 2005) and could therefore transmit epigenetic information throughout cell lineages. Moreover, neither a histone H4-K20-specific demethylase nor an H4 variant has been identified, making this modification potentially the most stable repressive methylation mark.

Chapter 5 ♦ Expression and distribution of Smyd1 and Prdm12 during myogenesis

* Experiments using the HA-Prdm12 construct have been performed by Ning Hu.

Introduction

Histone lysine methylation is currently one of the most widely studied chromatin modifications for its importance in gene regulation, development and cancer. The broad roles of these modifications have been largely identified by studying the enzymes that catalyze them, histone lysine methyltransferases (HKMTases). With the exception of Dot1, all HKMTases contain a SET/PR domain, which was originally identified through homology of Su-var39, Enhancer of Zeste and Trithorax proteins. Also common to most HKMTases is seemingly complex regulatory roles, catalyzing either single or multiple methylation states on one or more lysine residues.

A significant number of HKMTase family members have been implicated in human malignancy and it is therefore important to characterize their function (Huang 2002; Schneider et al. 2002; Kim et al. 2003; Kim and Huang 2003). RIZ1 was one of the first HKMTases shown to function as a tumor suppressor and several other RIZ1 related PR-domain containing proteins appear to have similar properties (Huang 2002). Another well known example of a cancer causing HKMTase is MLL, the most commonly disrupted gene in chromosomal rearrangements associated with leukemia (Daser and Rabbitts 2004). Although several other HKMTases are deregulated in cancer, their mechanisms of action on chromatin in many cases remain elusive. Nevertheless, the importance of understanding HKMTase function is highlighted by the progress that has been made from studying histone acetylation, which has led to the development of histone deacetylase inhibitors now used for chemotherapeutic purposes (Espino et al. 2005; Monneret 2005). Importantly,

both histone lysine acetylation and methylation patterns have the potential to serve as predictive diagnostic tools in cancer (Seligson et al. 2005).

Cancer often arises from deregulation of genes that control the cell cycle and differentiation and it is therefore not surprising that several HKMTases are also important for mammalian development. For instance, loss of *NSD2* may cause Wolf-Hirschorn syndrome, whereas its overexpression can lead to multiple myeloma (Stec et al. 1998; Malgeri et al. 2000). The importance of HKMTases in embryogenesis is apparent by the severe developmental defects resulting from targeted disruption of Ezh2, G9a, NSD1, Suv39H1/Suv39H2 or Mll (Yu et al. 1995; Peters et al. 2001; Tachibana et al. 2002; Erhardt et al. 2003; Rayasam et al. 2003). In addition, the expression profiles of HKMTases in mouse and human transcriptomes suggest that many of these enzymes have tissue-specific functions.

Global patterns of histone lysine methylation have been described during myogenesis (Biron et al. 2004) and although the enzymatic basis is unknown, some HKMTases have been implicated in muscle development. Smyd1 has emerged as an important regulator of myogenesis, as it is highly expressed in developing cardiac and skeletal muscle and its targeted loss results in heart malformations that are embryonic lethal (Gottlieb et al. 2002; Sims et al. 2002; Phan et al. 2005). Recently, Blimp1/Prdm1 was shown to be necessary for slow muscle development in zebrafish (Wilm and Solnica-Krezel 2005). A role for Ezh2 in skeletal myogenesis was also shown by its ability to regulate myogenic genes through MeK27-H3 (Caretti et al. 2004). The myogenic system therefore appears as an excellent model to examine the role of histone lysine methytion in cell differentiation. To

this end, we sought to identify and characterize HKMTases that potentially regulate myogenesis. By *in silico* analysis of mouse and human transcriptomes, we identified HKMTases with enriched expression in muscle, namely Prdm12 and Smyd1, and selected them for further molecular analysis. We obtained insight into the cellular distributions of Prdm12 and Smyd1 and their potential roles during muscle differentiation.

Results

In silico expression profiling of HKMTases

To determine the tissue distributions of HKMTases, we examined their expression profiles *in silico* using Unigene and SymAtlas transcriptome databases. Unigene contains expressed sequence tags (EST) from protein-coding nuclear genes from various tissues, whereas SymAtlas provides mRNA expression profiles from microarray analysis. Of 47 Set domain proteins listed in Entrez, 43 HKMTases had available expression profiles in Unigene's EST Profiles in 31 human tissues. We determined the mean tissue expression for each HKMTase and tabulated tissues that had expression 2-fold or 3-fold beyond the mean. For comparison, we performed a similar expression analysis using data available for 43 HKMTases in SymAtlas and reported only tissues that were shown in Unigene, with the exception of cardiac muscle where we included both whole heart and cardiomyocytes (Table 5-1). Human transcriptome data was used for comparison of both databases as it included a broader range of tissues with a greater number of HKMTases than currently available in the mouse, with the exception of Smyd1 where we used mouse data.

When comparing Unigene and SymAtlas data, there is significant disparity in the expression profiles of HKMTases (Table 5-1). Tissue elevations common to both databases are seen for only 7 HKMTases at 2-fold above mean levels and for only 2 HKMTases (Mll and Smyd1) at 3-fold above mean, and in no case is there complete concordance. To determine which database may be a more reliable expression prediction tool, we compared the expression profiles of genes with known involvement in muscle development, namely *MyoD1*, *Mef2C* and *Myf6* (Table 5-1). Consistent with their documented expression patterns (Berkes and Tapscott 2005), Unigene expression profiles of *MyoD1* and *Myf6* are highly specific for muscle and a significant elevation of *Mef2C* is seen in heart tissue. However, using SymAtlas, only the profile of *Myf6* is consistent with its known expression. In addition, the known tissue-specific expression profiles of HKMTases are more consistent with Unigene expression profiles. For instance, *Smyd1* expression is highly restricted to skeletal and cardiac muscle during embryogenesis and *PRDM8* is known to be elevated in the eye by RNA *in situ* hybridization (Mouse retina SAGE library). This suggests that Unigene expression data is a more reliable method for examining HKMTase expression profiles. Differences in reliability between these databases may be inherent to their methods of detection: Unigene expression is derived from expressed sequence tags (ESTs), whereas SymAtlas expression is obtained from Affymetrix hybridization arrays, which may give several false positives especially for genes such as HKMTases, which may have low expression. In addition, microarray tags are shorter sequences that may not have as much gene specificity. Unigene expression analysis revealed that 11/43 HKMTases have tissue-restricted distributions (17 tissues or

less) and significant tissue-specific enrichment (>3x mean), 22/43 HKMTases are broadly expressed with significant tissue-specific enrichment (>3x mean), and 10/43 HKMTases are broadly expressed without any tissue-specific enrichment (Table 5-1). This suggests that some HKMTases may have more general functions, whereas others have more specialized functions.

Prdm12 and Smyd1 in striated muscle development

To identify HKMTases with potential involvement in striated muscle development, we examined the expression profiles of 43 HKMTases for elevated expression in cardiac and skeletal muscle tissues in Unigene and SymAtlas transcriptomes. To this end, HKMTases were identified in both databases that exhibited expression levels at least >2x above mean in striated muscle (cardiac and skeletal). We identified 6 candidates by these criteria, namely, *C21orf18, EVI1 (PRDM3), PRDM4, PRDM5, PRDM12 and SMYD1*. *PRDM12* and *SMYD1* were selected for further analyses as they have highly elevated striated muscle expression (>3x mean) in both databases and are tissue-restricted in Unigene (Table 5-1, Figure 5-1 and 5-2).

Characterization of chromatin modifying enzymes in myogenesis

Expression dynamics of Smyd1 during skeletal muscle differentiation

We examined the dynamics of *Smyd1* expression in skeletal muscle differentiation by mining a gene array dataset from C2C12 myoblasts undergoing a 12-day developmental course. Tomczak et al. collected RNA daily for 2 days prior to differentiation induction and 10 days following differentiation induction at 2 day intervals (Tomczak et al. 2004).

For comparison, we displayed the expression of other HKMTases over the course of myogenesis (Figure 5-3 A). In proliferating myoblasts, *Smyd1* expression is relatively low and is comparable to other HKMTases with no predicted involvement in myogenesis. Upon induction of differentiation by serum withdrawal, *Smyd1* expression is dramatically elevated and persists throughout the 10-day differentiation time course. Importantly, *Ezh2* expression levels decline considerably upon serum withdrawal, consistent with its published expression in proliferating cells, which has been shown in myoblasts (Caretti et al. 2004). Furthermore, the expression of *Pr-Set7*, an mMeK20-H4-specific methylase (Xiao et al. 2005), also decreases upon differentiation, consistent with our results in chapters 3 and 4 showing a decrease in mMeK20-H4 upon induction of muscle differentiation. Consistent with published expression of *Smyd1*, the expression of this HKMTase would suggest it plays a role in skeletal muscle differentiation.

Overexpression analysis of Smyd1 in skeletal muscle differentiation

In some instances, the subnuclear localization of HKMTases is related to its function, as demonstrated by the distribution of SU(VAR)3-9 in pericentromeric heterochromatin domains enriched in tMeK9-H3 (Schotta et al. 2002). Even though the steady-state distribution of HKMTases may not reflect a carbon-copy of their chromatin substrates, localization may provide a good indication of their general chromatin associations. In the case of Smyd1, its general cellular localization has been shown (Gottlieb et al. 2002; Sims et al. 2002), but its subnuclear distribution has not been reported in detail. We therefore generated a YFP fusion protein with Smyd1 to further examine its localization and dynamics (Figure 5-4 A). YFP was fused to the carboxyl terminus of Smyd1 as this region

does not contain known functional domains. When expressed in C2C12 cells, the Smyd1-YFP expression plasmid produced a fusion protein of expected size by immunoblot analysis (Figure 5-4 A). In proliferating C2C12 myoblasts, Smyd1-YFP was distributed in both the nucleus and cytoplasm, consistent with the previously reported immunolocalization of endogenous Smyd1 (Gottlieb et al. 2002; Sims et al. 2002). In addition, an HA-tagged Smyd1 fusion protein displayed the same localization pattern as Smyd1-YFP (data not shown), suggesting the YFP fusion protein was not affecting the distribution of Smyd1.

To investigate the distribution of Smyd1, the localization of Smyd1-YFP was examined with respect to histone lysine methylation antibodies that demarcate distinct chromatin domains (shown in Chapters 3 and 4). The subnuclear distribution of Smyd1-YFP was excluded from pericentromeric and facultative heterochromatin domains as shown by an absence of overlap with DAPI-rich regions or tMeK20-H4 and mMeK20-H4, respectively (Figure 5-5 A-C). By comparison with tMeK4-H3, a mark of active euchromatin (Figure A-1, Appendices), Smyd1-YFP appeared to be enriched in euchromatic regions of the nucleus and this distribution persisted throughout C2C12 differentiation (Figure 5-5 D). The proximity of Smyd1 with tMeK4-H3 euchromatin domains may suggest its methylase specificity is similar to Smyd3, a known H4K3me2 and me3 HKMTase (Hamamoto et al. 2004). Consistent with this idea, Smyd1 is thought to catalyze H3-K4 methylation (Tan et al. 2006).

We attempted to uncover the methyl-lysine residue specificity of Smyd1 by examining a range of histone methyl-lysine derivatives in Smyd1-YFP transfected C2C12 cells.

Smyd1-YFP did not significantly colocalize with or alter the global levels of tMeK4-H3, tMeK9-H3, dMeK23-H3, d/tMeK36-H3, m/d/tMeK79-H3 or m/d/tMeK20-H4. Therefore, Smyd1 overexpression does not have a significant influence on the above histone lysine methylations, but it is possible that Smyd1 overexpression could influence a methylation not examined here or that overexpression of this protein does not significantly alter its methylation target. An extended analysis examining a larger panel of methyl-lysine residues in addition to Smyd1 knockdown experiments may reveal its methylase specificity *in vivo*.

Overexpression analysis of Prdm12 in skeletal muscle differentiation

As *in silico* expression analysis suggests Prdm12 has a role in myogenesis, we investigated its function in the context of histone methylation during muscle differentiation. To examine the subnuclear distribution of Prdm12, we generated a HA-tagged Prdm12 (HA-Prdm12) construct for immunofluorescent visualization using an anti-HA antibody. When overexpressed in C2C12 cells, HA-Prdm12 was diffusely distributed in the nucleus, excluded from DAPI-rich heterochromatin (Figure 5-6 A). Interestingly, HA-Prdm12 was generally concentrated in foci of variable sizes after 1 day of differentiation, but was more diffusely distributed after 6 days of differentiation. Because histone H4-K20 methylation undergoes change during muscle differentiation, we further examined the distribution of HA-Prdm12 with respect to these methyl-lysines by co-localization analysis, but we could not detect any significant co-localization between HA-Prdm12 and m/d/tMeK20-H4 (data not shown).

We also examined the influence of HA-Prdm12 on myotube formation and global levels of methylation. After 1 day of differentiation, HA-Prdm12 transfected cells did not appear grossly different than untransfected cells. Following 6 days of differentiation, HA-Prdm12 transfected cells were preferentially associated with multinucleated myotubes, although several untransfected cells were mononuclear in the culture (Figure 5-6 B). We examined the effects of HA-Prdm12 on muscle differentiation and d/tMeK20-H4, by determining whether HA-Prdm12 was associated with generally higher or lower levels of d/tMeK20 levels following differentiation. For this analysis, we separated the intensity profiles of d/tMeK20-H4 into low (lowest 33.3%), mid (middle 33.3%) and high (highest 33.3%) values, and looked at the intensities associated with HA-Prdm12 transfected cells versus all cells in the field (Figure 5-6 C). HA-Prdm12 transfected cells contained mostly mid-range intensities of dMeK20-H4 and higher levels of tMeK20-H4. This is shown by an association of 70% (n=30) of HA-Prdm12 nuclei with high intensity tMeK20-H4, a marker of differentiated myotubes, compared to only 12% (n=309) of all cells. Enrichments in d- and tMeK20-H4 in HA-Prdm12 transfected cells could be an indirect result of differentiation. Nevertheless, these results suggest that Prdm12 promotes muscle differentiation.

We also performed experiments with a flag-tagged Prdm12 expression construct (flag-Prdm12), which has a varied caroboxyl terminus with 65 amino acids replaced by 12 distinct amino acids followed by the Prdm12 3' UTR (Figure 5-4 B). This Prdm12 variant is generated from a canonical splicing sequence, although this sequence has not been validated to occur as an *in vivo* product. To determine its subnuclear distribution, we

colocalized flag-Prdm12 (Figure 5-7, Table 5-2) with m/d/tMeK20-H4. flag-Prdm12 was distributed in ring-shaped domains in non-heterochromatic regions of the nucleus. We observed appreciable colocalization of flag-Prdm12 with dMeK20-H4 but not with mMeK20-H4 or tMeK20-H4 (Table 5-2, Figure 5-7). In addition, dMeK20-H4 was concentrated in Prdm12 domains and had an apparent overall enrichment in Prdm12-transfected nuclei (Figure 5-7 B). dMeK20-H4-Prdm12 colocalization and the global elevation of dMeK20-H4 was seen in C2C12 cells at day 1 to day 3 and primary limb bud cultures in ~ 60-80% of Prdm12–transfected cells. As dMeK20-H4 did not colocalize with cytoplasmic flag-Prdm12, and there is no significant sequence similarity between the H4K20 flanking sequence and Prdm12 or the flag epitope tag, we do not expect the dMeK20-H4 antibody to be cross-reacting with a methylated flag-Prdm12 protein, however, this possibility could be ruled out by performing a peptide competition assay. Prdm12-transfected cells also had an overall decrease in mMeK20-H4 but not tMeK20-H4, and this depletion did not appear to be attributed to a loss of facultative heterochromatin aggregates. Importantly, although the flag-Prdm12 construct appears to mislocalize in the nucleus in comparison with the full length protein, this data could suggest that the SET domain of Prdm12 mediates the catalysis of dMeK20-H4. However, further analysis would be needed to confirm such a function.

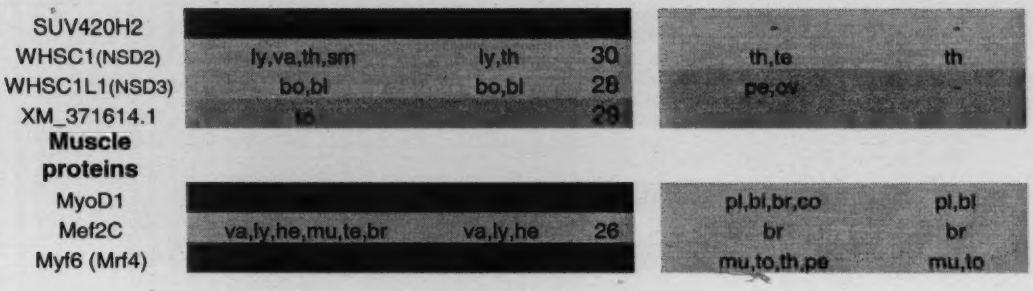
We further examined the influence of flag-Prdm12 overexpression on skeletal muscle differentiation in C2C12 cells and primary limb bud cultures. In both C2C12 and primary cultures, flag-Prdm12 transfected cells did not develop into multinucleated myotubes after 7 days and therefore appeared to repress skeletal muscle differentiation. In addition, the

levels of Myod and myogenin were either absent or below antibody detection in flag-Prdm12 transfected C2C12 and primary culture cells (Table 5-3, Figure 5-8). This data suggests that flag-Prdm12 perturbs the ability of skeletal myoblasts to differentiate, however it remains to be found whether Prdm12 functions to repress myogenesis or if this construct interferes with the endogenous activity of this protein.

Table 5- 1: Summary of *in silico* gene expression analysis from 45 HKMTases. Human expression data obtained from Unigene and SymAtlas was analyzed for 45 HKMTases (with exception of Smyd1 where only mouse expression was available in SymAtlas). For comparison, tissue expression in SymAtlas is reported for the 31 tissues shown in Unigene, with the exception of cardiac muscle where both cardiac myocytes and whole heart expression is reported from SymAtlas data. HKMTase tissue expression that is 2x or 3x above the mean expression of all tissues is indicated. HKMTases with expression limited to 17/31 tissues or less are highlighted in red. HKMTase expression was found in all tissues in SymAtlas, but was restricted for some HKMTases in Unigene. HKMTases with tissue-specific expression 3x above the mean are shown in light blue and those without tissue-specific expression 3x above the mean are indicated in green. Note that all tissue-restricted HKMTases (red) were also expressed with significant elevation (>3x above mean) for a subset of tissues. Gray shading indicates HKMTases where no expression data was available. Yellow highlighting indicates HKMTases with elevated striated muscle expression (heart, muscle or cardiac myocytes) found in both Unigene and SymAtlas. SymAtlas tissues were scored as positive if the error bars surpassed 2x or 3x mean. ba, bladder; bl, blood; bo, bone; bom, bone marrow; ce, cervix; co, colon; ey, eye; he, heart; hem, cardiac myocyte; ki, kidney; la, larynx; li, liver; lu, lung; ly, lymph node; ma, mammary gland; mu, muscle (skeletal); ov, ovary; pa, pancreas; pe, peripheral nervous system (obtained from trigeminal ganglion and spinocervical ganglion in SymAtlas); pl,

placenta; pr, prostate; sk, skin; sm, small intestine; so, soft tissue; sp, spleen; st, stomach; to, tongue; te, testis; th, thymus; ut, uterus; va, vascular.

	Unigene			SymAtlas	
HKMTases	*>2x mean*	*>3x mean*	*all*	*> 2x mean*	*>3x mean*
ASH1L	ma,th	th	29	-	-
BAT8 (G9a)	pr	-	29	th	-
C14orf154	ki, bo		29		
C21orf18	sp, pe, lu, he, pl, co	sp, pe	21	te,pr,hem	te
Dot1L	ly,te,sp,lu,bl	ly	23	-	-
EHMT1	ly,pe,bl,th	ly	24	mu	
EVI1 (PRDM3)	bl, mu,st,ut, ki, co	bl, mu, st	26	ki,pa,va,lu,ut,hem	pa,t,hem
EZH1	sp,ey,pr,so		28	ut	-
EZH2	la,sm,ly,te		28	th,te,hem	th,te,hem
MLL	ly,th,bl,la	ly,th	30	ly,th,ut	ly,th
MLL2	ba,bl,la,th,sm	ba,bl, la	28	li	-
MLL3	la,bl,pl	la	28	bl,pl,th	bl,pl
MLL4	th	-	26	co,th	co
MLL5	la,ma,to		28	pl,ut,bl	bl
NSD1	la,bl,ut,ma	la	24	-	-
PRDM1				li,pe,mu	li
PRDM2	bl,ly,sp	bl	25	bl,br,pe	br
PRDM4	bo,ki,ma,pa,he	bo	21	va,hem	-
PRDM5	pe,la,to,mu	pe,la,to	22	he	-
PRDM6				-	-
PRDM8				ov	-
PRDM9				mu,li,pe,hem	-
PRDM10	th,bom	th,bom	19	co,pr,hem	-
PRDM11					
PRDM12				pa,he,hem,li,pl,br	hem,li
PRDM13				pe,hem	-
PRDM14				-	-
PRDM15				-	-
PRDM16					
SET7	th,ly,pe	-	28	va,pr	pr
SET8	la	la	23	hem,pr	hem,pr
SETDB1	ba,so,to,sp	ba,so,to	27	te,bl	-
SETDB2					
SMYD1				mu,he,to,bo,bom,pe	mu,he,to,bo
SMYD2	ki,br	-	26	pe	pe
SMYD3	pe,st,bo,li,bom	pe	25	br,lu	br
SMYD4	sm,li,bom,he,co,pa,to		18	-	-
SMYD5	ki,so,te,mu	ki,so	21	pe	-
SUV39H1	th,sm,ce,ov	th,sm,ce,ov	24	co,li	co,li
SUV39H2	ce,sm,mu,ki	ce	21	-	-
SUV420H1	ma,la,pe	ma	26	th,co	-


SUV420H2
WHSC1(NSD2)
ly,va,th,sm
ly,th
30
th,te
th
WHSC1L1(NSD3)
bo,bl
bo,bl
28
pe,ov
-
XM_371614.1
te
28
Muscle proteins
MyoD1
pl,bl,br,co
pl,bl
Mef2C
va,ly,he,mu,te,br
va,ly,he
26
br
br
Myf6 (Mrf4)
mu,to,th,pe
mu,to

Transfected Cells	Antibody	Colocalization	Association with increased or decreased levels
C2C12 (day 1)	mMeK20-H4	-	decreased
C2C12 (day 1)	dMeK20-H4	+	increased
C2C12 (day 1)	tMeK20-H4	-	-
C2C12 (day 1)	tMeK4-H3	-	-
C2C12 (day 3)	mMeK20-H4	-	decreased
C2C12 (day 3)	dMeK20-H4	+	increased
C2C12 (day 3)	tMeK20-H4	-	-
C2C12 (day 3)	tMeK4-H3	-	-
C2C12 (day 7)	mMeK20-H4	-	decreased
C2C12 (day 7)	dMeK20-H4	+	increased
C2C12 (day 7)	tMeK20-H4	-	-
C2C12 (day 7)	tMeK4-H3	-	-
RH-30	mMeK20-H4	-	-
RH-30	dMeK20-H4	+	increased
RH-30	tMeK20-H4	-	-
Primary cultures (day 1)	mMeK20-H4	-	decreased
Primary cultures (day 1)	dMeK20-H4	+	increased
Primary cultures (day 1)	tMeK20-H4	-	-
Primary cultures (day 3)	mMeK20-H4	-	decreased
Primary cultures (day 3)	dMeK20-H4	+	increased
Primary cultures (day 3)	tMeK20-H4	-	-
Primary cultures (day 7)	mMeK20-H4	-	decreased
Primary cultures (day 7)	dMeK20-H4	+	increased
Primary cultures (day 7)	tMeK20-H4	-	-

Table 5- 2: Immunofluorescence analysis of flag-Prdm12 overexpression on histone lysine methylations.

Cells were transfected with a flag-Prdm12 expression construct and colocalized with anti-Prdm12 and histone methyl lysine antibodies. Colocalization was determined at 100x magnification and overall elevations or depletions in genome-wide methylation of transfected cells were determined at 20x.

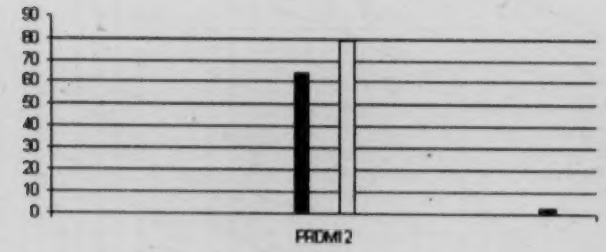
Transfected Cells	Antibody	Colocalization	Association with increased or decreased levels
C2C12 (d 1)	Myod	-	decreased
C2C12 (d 1)	myogenin	-	decreased
C2C12 (d 3)	Myod	-	decreased *
C2C12 (d 3)	myogenin	-	decreased *
C2C12 (d 7)	Myod	-	decreased
C2C12 (d 7)	myogenin	-	decreased
Primary cultures (d 1)	Pax3	-	-
Primary cultures (d 1)	Pax7	-	-
Primary cultures (d 1)	Myod	-	decreased (12/12)
Primary cultures (d 1)	Myogenin	-	decreased (12/12)
Primary cultures (d 1)	Myf-5	-	-
Primary cultures (d 1)	BrdU	-	1/14 transfectants BrdU positive
Primary cultures (d 3)	Pax3	-	-
Primary cultures (d 3)	Pax7	-	-
Primary cultures (d 3)	Myod	-	decreased (63/63)
Primary cultures (d 3)	Myogenin	-	decreased (50/50)
Primary cultures (d 3)	Myf-5	-	-
Primary cultures (d 3)	BrdU	-	0/16 transfectants BrdU positive
Primary cultures (d 7)	Pax3	-	-
Primary cultures (d 7)	Pax7	-	-
Primary cultures (d 7)	Myod	-	-
Primary cultures (d 7)	Myogenin	-	-
Primary cultures (d 7)	Myf-5	-	-
Primary cultures (d 7)	BrdU	-	-

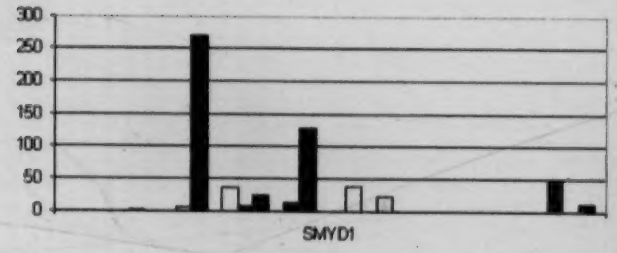
Table 5- 3: Immunofluorescence analysis of flag-Prdm12 overexpression on muscle cell markers.

Cells were transfected with a flag-Prdm12 expression construct and colocalized with either anti-flag or anti-Prdm12 and skeletal muscle cell antibodies. To track cell cycle changes occuring in primary cultures, BrdU was added to cultures for 3 hrs prior to fixation for immunofluorescence staining. Colocalization was examined at 100x magnification and overall elevations or decreases were examined at 20x. * 103/103 Prdm12-transfected C2C12 cells were Myod negative and 336/337 were Myogenin negative, in cultures staining with ~ 20-30% Myod and ~5-10% Myogenin positive.

Figure 5- 1: Unigene expression profile of Smyd1 and Prdm12.

The expression of Prdm12 and Smyd1 are shown in transcripts per million (tpm) in 31 human tissues in addition overall levels in embryo, juvenile and adult tissues. Prdm12 expression in muscle is above our designated cutoff of 3x above mean, as is the expression of Smyd1 in both heart and muscle.


90
80
70
60
50
40
30
20
10
0
PRDM12


300
250
200
150
100
50
0
SMYD1

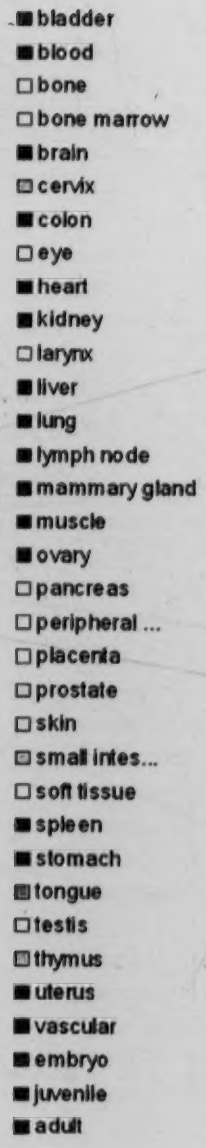

bladder
blood
bone
bone marrow
brain
cervix
colon
eye
heart
kidney
larynx
liver
lung
lymph node
mammary gland
muscle
ovary
pancreas
peripheral ...
placenta
prostate
skin
small intes...
soft tissue
spleen
stomach
tongue
testis
thymus
uterus
vascular
embryo
juvenile
adult

Figure 5- 2: SymAtlas expression profile of Smyd1 and Prdm12.

(A) The expression profile of Smyd1 obtained from a large-scale analysis of the mouse transcriptome (GNF1M) comparing the expression of 61 physiologically normal tissues interrogated with 36, 182 probe sets. In this dataset, significant expression of Smyd1 is noted in cardiac and skeletal muscles with values 10 fold above the mean expression of all tissues. (B) The expression profile of Prdm12 obtained from a large-scale analysis of the human transcriptome (GNF1H) comparing the expression of 79 physiologically normal tissues interrogated with 44, 775 probe sets. The expression of Prdm12 in cardiac muscle is shown to be 10 fold above the mean expression of all tissues.

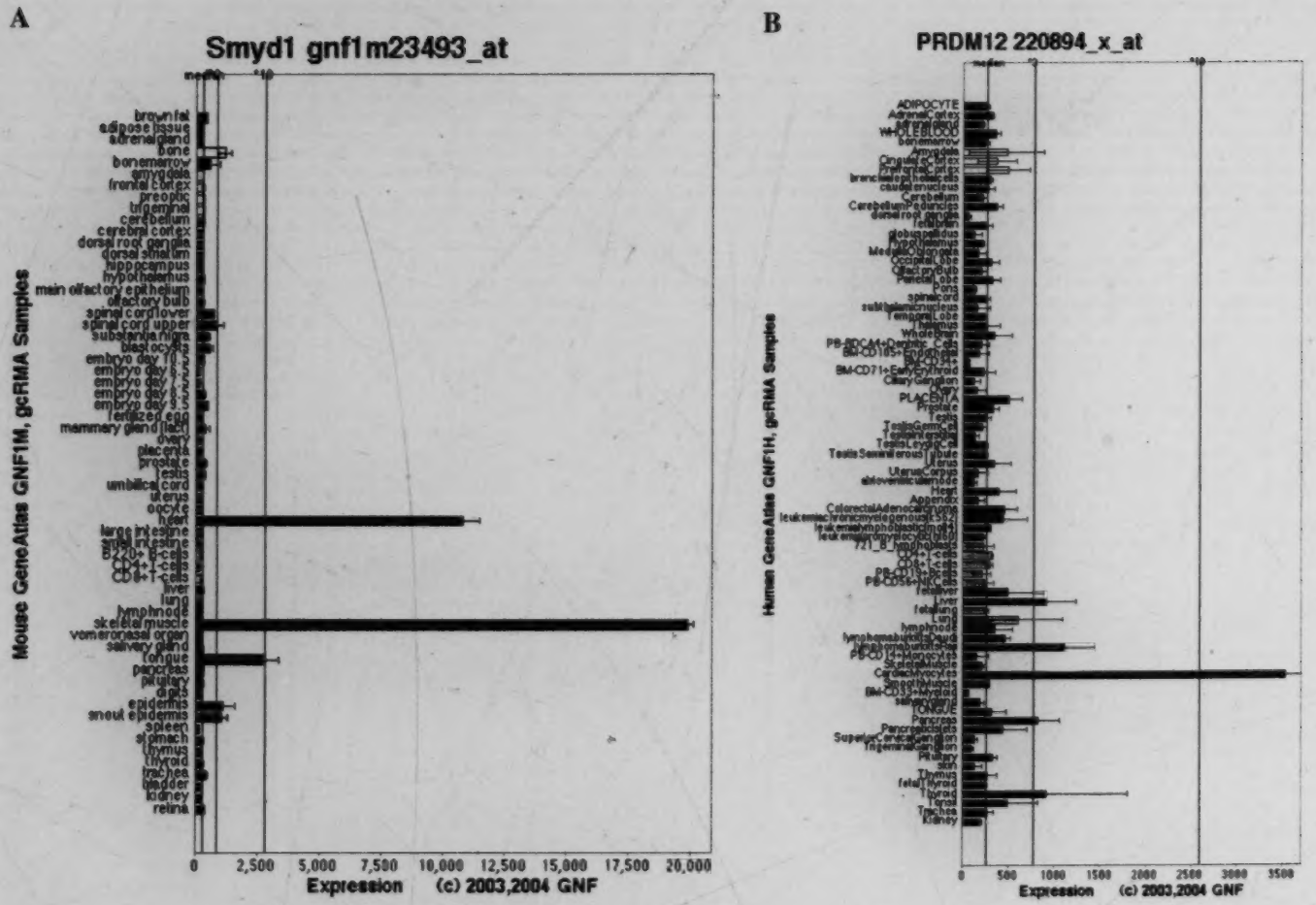

A
Smyd1 gnf1m23493_at
Mouse GeneAtlas GNF1M, gcRMA Samples
brown fat
adipose tissue
adrenal gland
bone
bonemarrow
amygdala
frontal cortex
preoptic
trigeminal
cerebellum
cerebral cortex
dorsal root ganglia
dorsal striatum
hippocampus
hypothalamus
main olfactory epithelium
olfactory bulb
spinal cord lower
spinal cord upper
subtantia nigra
blastocyst
embryo day 10.5
embryo day 11.5
embryo day 12.5
embryo day 13.5
embryo day 9.5
fertilized egg
mammary gland lact
ovary
placenta
prostate
testis
umbilical cord
uterus
oocyte
heart
large intestine
small intestine
B220+ B-cells
CD4+ T-cells
CD8+ T-cells
liver
lung
lymphnode
skeletal muscle
vomeronasal organ
salivary gland
tongue
pancreas
pituitary
aorta
epidermis
snout epidermis
spleen
stomach
thymus
thyroid
trachea
bladder
kidney
retina
0
2,500
5,000
7,500
10,000
12,500
15,000
17,500
20,000
Expression
(c) 2003,2004 GNF
B
PRDM12 220894_x_at
Human GeneAtlas GNF1H, gcRMA Samples
ADIPOCYTE
AdrenalCortex
WHOLEBLOOD
bonemarrow
Amygdala
CingulateCortex
FrontalCortex
caudatenucleus
Cerebellum
Cerebellum Peduncles
dorsalrootganglia
Fetalbrain
Globuspallidus
Hypothalamus
Occipital Lobe
OlfactoryBulb
Parietal Lobe
Pons
spinalcord
subNigranucleus
Temporal Lobe
Thalamus
WholeBrain
PB-BDCA4+Dendritic Cells
BM-CD105+Endothelial
BM-CD71+EarlyErythroid
CiliaryGanglion
Ovary
PLACENTA
Prostate
Cells
TestisGermCell
TestisInterstitial
TestisLeydigCell
TestisSeminiferous Tubule
Uterus
UterusCorpus
chloroleukemiccelline
Heart
Appendix
ColorectalAdenocarcinoma
leukemicchronicmyeloidgenous
leukemiclymphoblasticmyeloid
leukemiclymphocyticchronic
721_B_lymphoblast
CD4+ Cells
CD8+ Cells
PB-CD19+B-Cells
PB-CD56+NK-Cells
Fetalliver
Liver
fetallung
Lung
lymphnode
lymphomaburkittsraji
PB-CD14+Monocytes
SkeletalMuscle
CardiacMyocytes
BM-CD33+Myeloid
salivarygland
TONGUE
Pancreas
Pancreaticislets
SuperiorCervicalGanglion
trigeminalganglion
Pituitary
skin
Thymus
fetalThyroid
Thyroid
Tonsil
Trachea
Kidney
0
500
1000
1500
2000
2500
3000
3500
Expression
(c) 2003,2004 GNF

Figure 5- 3: HKMTase expression during C2C12 skeletal muscle differentiation.

(A) The expression profile of *Smyd1* in relation to other HKMTases, chromatin modifiers and muscle-specific transcription factors from gene array data (http://www.chb-genomics.org/beggslab/) are shown in arbitrary units over the course of C2C12 muscle differentiation in 2 day increments. Cells at day -1 and -2 are in proliferation media and attain 100% confluence at day 0 when they are induced to differentiate for 10 days by serum withdrawal (Tomczak et al. 2004). (B) RT-PCR of *Smyd1* (skm-Bop) and *Prdm12* (lane 3) using C2C12 RNA, amplifying products of expected size. Prdm12 RT-PCR in lane 1 was performed with RNA of poor quality (frozen > 1 year) and the reaction in lane 2 was performed with an RNA concentration 4-fold higher than in lane 3.

A
1000
100
10
1
0.1
0.01
Smyd1
Mef2c
Myf6
Myogenin
Ezh1
Rb1
Mef2a
Hdac5
PRDM1
MyoD1
Suv39h2
Ezh2
Smyd2
E2f1
Myf5
Hdac1
E2f5
PRSET-7
Smyd5
SETDB1
SETDB1
Smyd5
PRSET-7
E2f5
Hdac1
Myf5
E2f1
Smyd2
Ezh2
Suv39h2
MyoD1
PRDM1
Hdac5
Mef2a
Rb1
Ezh1
Myogenin
Myf6
Mef2c
Smyd1
B
1 2 3
Smyd1
PRDM12

Figure 5- 4: Schematic of PRDM12 and SMYD1 domain structure and fusion proteins used for their functional analysis.

(A) Full length human SMYD1 containing a split S-ET domain and a MYND domain, was fused to eYFP for immunolocalization studies. Previously reported protein interactions are indicated above. SMYD1-YFP produces a protein of expected length as determined by immunoblot analysis (right) showing a blot with whole cell extracts from C2C12 cells transfected with 1) pEYFPN.1 empty vector, 2) SMYD1-EYFP following 24 hrs of differentiation, 3) SMYD1-EYFP following 72 hours of differentiation, and 4) HA-tagged SMYD1 probed with anti-GFP. (B) Human PRDM12 was fused to an HA-tag at the amino terminus. Flag-Prdm12 contains a flag tag at the start of the protein and the construct is truncated at the last 65 amino acids, resulting in the loss of zinc finger domains (Znf) with the introduction of 12 new amino acids as a result of a splicing event.

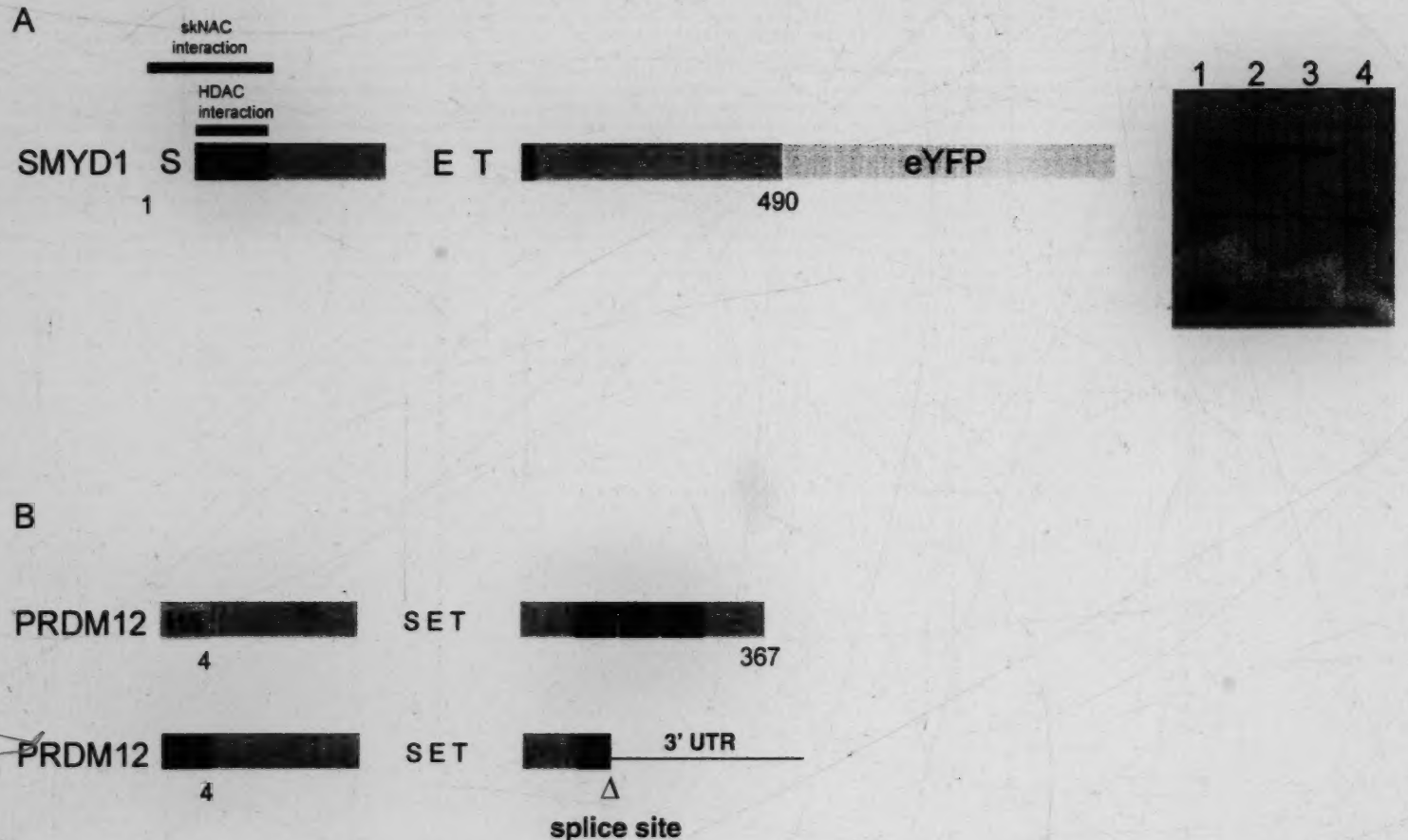
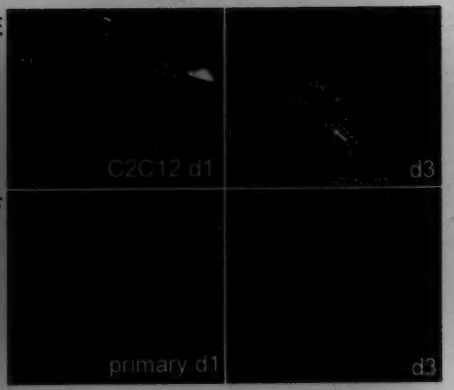

A
skNAC interaction
HDAC interaction
SMYD1 S
E T
eYFP
1
490
1 2 3 4
B
PRDM12
SET
4
367
PRDM12
SET
3' UTR
4
Δ
splice site

Figure 5- 5: Functional analysis of Smyd1 in C2C12 cultures and primary cells

(A-D) The subnuclear distribution of Smyd1-YFP (center panels) is shown with respect to (A) DAPI, (B) tMeK20-H4, (C) mMeK20-H4 and (D) tMeK4-H3. Images are merged in right hand panels showing Smyd1-YFP in green with respect to DAPI/histone antibody in red. (E, F) Smyd1-YFP transfected (E) C2C12 cells and (F) primary limb bud cultures following 1 (left panels) or 3 (right panels) days of differentiation. (G) Percentage of cells with myotube morphology in cells transfected with eGFP, Smyd1-YFP and Flag-Prdm12. following 1, 3 or 7 days of differentiation.



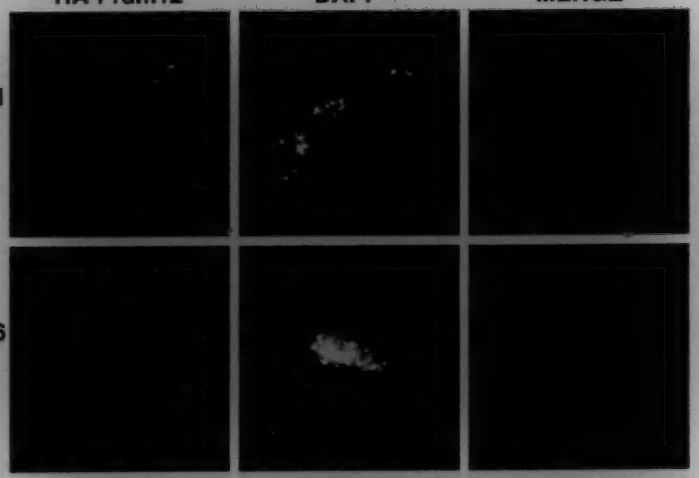
G

	Pecentage of cells with myotube morphology		
Days post-transfection	eGFP	Smyd1-YFP	Flag-PRDM12
day 1	5.2 (n=230)	0 (n=92)	0 (n=47)
day 3	5.8 (n=104)	0 (n=1270)	0 (n=286)
day 7	23.3 (n=107)	8.1 (n=62)	0 (n=16)

Figure 5- 6: Overexpression analysis of HA-Prdm12 in C2C12 cells.

(A) Subnuclear distribution of HA-Prdm12 in C2C12 cells following 1 or 6 days of differentiation. (B) Distribution of HA-Prdm12 with respect to dMeK20-H4 (top panels) or tMeK20-H4 (bottom panels). The intensity profiles of d/tMeK20-H4 are shown in the rightmost panels (red= highest 33.3%, blue=middle 33.3%, green=lowest 33.3%). (C) The scored intensity profiles of d/tMeK20-H4 in HA-Prdm12 transfected cells compared to all cells.

A


HA-Prdm12
DAPI
MERGE
DAY 1
DAY 6

B

HA-Prdm12
dMeK20-H4
intensity profile
Prdm12-HA
tMeK20-H4
intensity profile

C

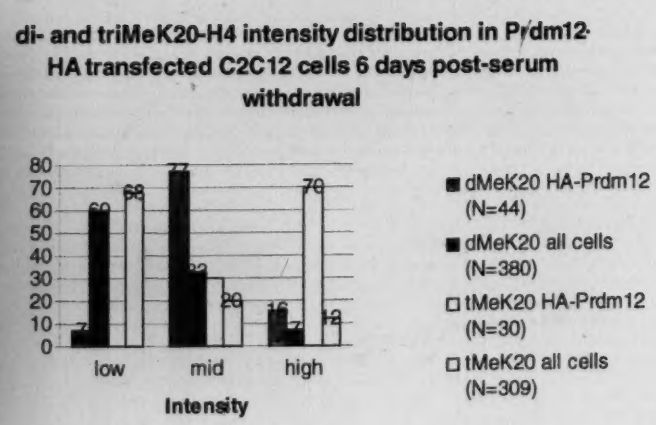

di- and triMeK20-H4 intensity distribution in Prdm12-HA transfected C2C12 cells 6 days post-serum withdrawal
80
70
60
50
40
30
20
10
0
low
mid
high
Intensity
dMeK20 HA-Prdm12 (N=44)
dMeK20 all cells (N=380)
tMeK20 HA-Prdm12 (N=30)
tMeK20 all cells (N=309)

Figure 5- 7: Overexpression analysis of flag-Prdm12 in C2C12 cultures and primary cells.

(A-C) Subnuclear distribution of flag-Prdm12 (left panels) with respect to (A) mMeK20-H4, (A-crop) magnified view, (B) dMeK20-H4, (B-crop) magnified view and (C) tMeK20-H4, (C-crop) magnified view (center panels). Merged images are shown in right hand panels, with flag-Prdm12 colored in red and MeK20-H4 shown in green.

A
DAPI
H4K20me1
PRDM12
MERGE
A-crop
B
DAPI
H4K20me2
PRDM12
MERGE
B-crop
C
DAPI
H4K20me3
PRDM12
MERGE
C-crop

Figure 5- 8: Localization of flag-Prdm12 and markers of muscle differentiation.

(A) Localization of flag-Prdm12 (red) with respect to DAPI (Blue) and MyoD (Green) or (B) myogenin (red), following 3 days of differentiation in C2C12 cultures. (C) Colocalization of flag-Prdm12 (green) with respect to DAPI (blue) and MyoD (red) or (D) myogenin (red), following 3 days of differentiation in mouse primary limb bud cultures.

A
MyoD
PRDM12
DAPI
MERGE
C2C12-d3
B
Myogenin
C2C12-d3
C
MyoD
primaries-d3
D
Myogenin
primaries-d3

Discussion

Tissue specific roles for HKMTases

Histone lysine methylation is an attractive candidate for regulating cell memory – the ability of cell to transfer regulatory knowledge to its daughter – and may have the potential to inscribe information necessary for the determination and maintenance of cell identity. If such a role were mediated by this modification, one could hypothesize that there are cell-type specific HKMTases present to propagate and maintain a particular methylation profile unique to a cell's identity. Alternatively, specificity could be conferred by other proteins that recruit the action of HKMTases. Our *in silico* gene expression analysis of 45 HMTases suggests that some HKMTases may provide lineage-specific epigenetic regulation by virtue of their tissue restricted expression – an idea supported by current literature (Lin and Dent 2006). Consistent with previous studies, our *in silico* profiles would suggest that Suv39h1/2, Ezh2 and Bat8 (G9a) have regulatory functions in a broad range of cell types, whereas Smyd1 may have a more lineage-restricted function in striated muscle. Interestingly, it appears that several *PRDM* genes have lineage-restricted roles. It is tempting to speculate that a common feature of PRDM proteins is to provide tissue-restricted histone methylation marks. However, from a mechanistic standpoint, little is known about most PRDM proteins at present.

With recent developments in high-throughput technologies, *in silico* profiling of transcriptomes offers a useful method for predicting gene expression (Yamane et al. 2006). Although this approach may not always correlate with *in vivo* expression analyses, an

association is present for some genes. In our analysis, *in vivo* to *in silico* correlations can be seen with Unigene expression profiles - *Ezh2* is thought to have a general role in cell proliferation and *Smyd1* appears to have a more restricted function in striated muscle, as suggested by *in vivo* work. In the case of Prdm12, which does not have *in vivo* expression data published, we expect that the *in silico* data indicating muscle expression is likely to be representative due to its elevated expression in relation to other tissues, similar to *Smyd1* expression.

HKMTase regulation of myogenesis

Suv39h and Ezh2 HKMTases have been examined in the context of skeletal muscle differentiation, but their influence on myogenesis likely underlies general growth regulatory functions applicable to many cell types. Cell cycle exit during muscle differentiation has been shown to be dependent on Suv39h-mediated repression of S-phase genes (Ait-Si-Ali et al. 2004). In addition, depletion of *Suv39h1/2* in myoblasts results in an inhibition of muscle differentiation markers, although this is likely due to an indirect effect of deregulated silencing of S-phase genes. In support of a more general role for Suv39h in differentiation, *Suv39h* dn (see Chapter 1) mice have global growth defects (Peters et al. 2001) and *Suv39h1/2* appear to be broadly expressed in human tissues (Table 5-1). The role of tMeK9-H3 in cell differentiation may be more gene-centric in regulating the cell cycle, suggested by our observations that tMeK9-H3 levels appear static over the course of skeletal muscle differentiation (Chapters 3 and 4). In addition, this modification appears to play a specific role in mitosis, which could be unique in muscle cells. Ezh2 has been shown to regulate myogenesis through tMeK27-H3-mediated repression of muscle

differentiation genes (Caretti et al. 2004) but it also appears to have a more general role in mediating proliferation of numerous cell types and in mediating development of ES cells (Boyer et al. 2006; Lee et al. 2006). This is apparent from the proliferation defects seen in *Ezh2*-deficient mice (O'Carroll et al. 2001), and its widespread distribution in various human tissues (Figure 4-1) and mouse mid-gestation embryos (Caretti et al. 2004). Of note, both tMeK27-H3 and mMeK20-H4 are thought to be enriched in the X_i and our observations (Chapters 3 and 4) indicate that mMeK20-H4 is generally elevated in proliferating cells, raising the possibility that an X_i mediated event may be involved in muscle differentiation. Therefore, it appears that tMeK9-H3 serves to repress the cell cycle, whereas tMeK27-H3 has a general role in repressing differentiation, and neither Suv39h1/h2 nor Ezh2 appear to distinctly specify or maintain myogenic identity.

Smyd1 in muscle development

Smyd1 (*m-Bop*) was first identified as a gene highly expressed in cytotoxic T-cells and adult striated muscle (Hwang and Gottlieb 1995), but was subsequently described as a gene important in heart development (Gottlieb et al. 2002). *SMYD1* has clinical importance as its RNA levels increase in end stage heart disease, suggesting that alterations in *SMYD1* activity may contribute to cardiac pathogenesis (Borlak and Thum 2003). In examining the role of *Smyd1* found it was found to be expressed at ~ E8 in mouse cardiac myocytes and at E12.5 in the myotome, and its targeted loss resulted in a severely hypomorphic heart similar to *Mef2C -/-* mice (Gottlieb et al. 2002). In addition, cardiomyocytes from *Smyd1*-deficient mice have decreased expression of the *Hand2* transcription factor (heart and neural crest derivatives expressed transcript 2), which disrupts cardiac differentiation.

Recently, it was shown that *Smyd1* is a direct target of Mef2C in a transcriptional cascade required for heart development (Phan et al. 2005), explaining the similarity between *Smyd1* and *Mef2C* nulls. Based on microarray expression data, it appears that Smyd1 and Mef2C are similary expressed during myogenesis (Figure 5-3). In skeletal myoblasts, *Smyd1* expression is elevated upon cell-cycle exit and its levels persist throughout differentiation (Sims et al. 2002; Tomczak et al. 2004). In addition, *Smyd1* appears to be regulated downstream of Myod (Blais et al. 2005) and associates with the transcriptional activator skNAC to promote muscle differentiation (Sims et al. 2002). Although Smyd1 is clearly important for cardiac and skeletal myogenesis, its molecular mechanism of action remains elusive. Importantly, Smyd1 contains a split SET domain (S-ET) and a MYND domain, which has been shown to mediate HDAC-dependent transcriptional repression (Figure 5-5) and is thought to catalyse methylation of histone H3-K4 (Tan et al. 2006). Therefore Smyd1 may promote and/or maintain muscle differentiation through a gene regulation mechanism that involves histone acetylation and methylation. Our analysis of Smyd1 thus far is consistent with published studies but further investigation is required to understand its role in the context of histone methylation.

Several lines of evidence implicate Smyd1 as a likely candidate for the distinct chromatin dynamics involved in muscle differentiation. Our *in silico* analysis revealed that *Smyd1* expression is tissue-restricted and significantly elevated in skeletal and cardiac muscle. Importantly, *Smyd1* levels persist during differentiation of C2C12 cells (Figure 5-3), which is consistent with the progressive histone methyl-lysine dynamics in this cell line and in embryonic muscle cultures (Chapter 3). As shown by our results in Chapter 3 and

the work of others, its expression throughout the cardiac system and the myotome also parallels the embryonic dynamics of tMeK20-H4 (Gottlieb et al. 2002). Although our subnuclear localization data would not suggest Smyd1 is involved in regulating tMeK20-H4 levels during embryogenesis, its expression on a temporal and tissue level coincides with global chromatin changes occurring in myogenesis. However, with regard to the function of Smyd1, it is difficult to interpret the how its overexpression in embryonic muscle cultures could prevent differentiation. One possibility is that overexpression of Smyd1 causes non-specific chromatin changes in addition to its normal function in a way that disrupts differentiation genes. Smyd1-YFP may not function normally, and could compete for endogenous Smyd1 activity. Performing targeted *Smyd1* knockdown experiments in muscle cells could perhaps better address the function of this HKMTase.

By examining the effects of Smyd1-YFP overexpression on C2C12 differentiation, we obtained preliminary results that should be elaborated for a more functional understanding of Smyd1. A previous report showed an apparent re-localization of endogenous nuclear Smyd1 to the cytoplasm in C2C12 cells following 4 days of differentiation, suggesting Smyd1 nuclear levels were either absent or very low in differentiated myotubes (Sims et al. 2002). In contrast, we observed persistent nuclear and cytoplasmic localization of Smyd1-YFP in differentiated myotubes, 10 T1/2 cells and primary limb bud cultures. This discrepancy may be explained by the different fixation times used for immunofluorescence of myoblasts (10 mins) versus myotubes (20 mins) reported by Sims et al. (Sims et al. 2002), a higher range of detection with our overexpressed fluorescent protein, or impaired function of Smyd1 by the YFP fusion protein. In C2C12 cells, Smyd1-YFP did not appear

to influence the generation of multinucleated myotubes (see Figure 5-5 E). In addition, overexpression of Smyd1 in 10T1/2 cells did not generate myotubes (data not shown), suggesting that Smyd1 does not have the capacity to induce myogenesis like the bHLH myogenic regulatory factors (MRFs). However, in mouse primary limb bud cultures, at least 1260 Smyd1-YFP positive cells did not differentiate into multinucleated myotubes following 3 days of differentiation in contrast to ~ 5% of GFP-transfected positive cells that were myotubes (Figure 5-5 G). By day 7, some Smyd1-YFP positive myotubes were seen, but was proportionately lower than cells transfected with GFP only. Consistent with a reduced number of Smyd1-YFP positive cells undergoing differentiation in primary cells, a small percentage of cells were positive for myogenin. These results suggest that overexpression of Smyd1 perturbs the progression of muscle differentiation.

Our localization analysis suggests Smyd1 in enriched in euchromatic domains by its association with tMeK4-H3-rich regions and may have a role in the cytoplasm. Although absolute co-localization was not observed between Smyd1-YFP and tMeK4-H3, their close association represents a preference of Smyd1 for euchromatic regions, consistent with the idea that this HKMTase catalyses histone H3-K4 methylation (Tan et al. 2006). This distribution could be related to the high level of similarity between Smyd1 and Smyd2, a dMeK36-H3 HKMTase and Smyd3, a d/tMeK4-H3 HKMTase, as these methyl-lysine modifications are distributed in DAPI-depleted regions of the nucleus (Brown et al. 2006). In support of similar functions for Smyd1 and Smyd3, depletion of these proteins appears to result in apoptosis in independent cell contexts (Gottlieb et al. 2002; Hamamoto et al. 2004). However, the expression of *Smyd2* in the embryonic heart (Brown et al. 2006)

suggests a possibly related developmental function between this HKMTase and Smyd1. Further dissection of Smyd1 function in comparison to Smyd2 and Smyd3 may facilitate its characterization. The persisting levels of Smyd1 in the cytoplasm shown here and reported by others, which is similar to Smyd2 and Smyd3, may also be of significance to its function. Recently, Ezh2 was shown to have a role in cytoplasmic actin polymerization in addition to its nuclear HKMTase function (Yamane et al. 2006). In Chapter 4, we have shown a striking redistribution of the m/tMeK79-H3 antibodies to the cytoplasm upon differentiation, which we could speculate to represent the signature of protein lysine methylation by a cytoplasmic HKMTase. It will be interesting to see whether future studies reveal dual cytoplasmic and nuclear functions for other HKMTases.

A potential role for Prdm12 in myogenesis

PRDM12 is part of the PR-domain-containing zinc-finger family, whose members have been shown to function as tumor suppressors (Reid and Nacheva 2004). The PR domain is closely related to the SET domain and has also been shown to catalyze histone lysine methylation (Rea et al. 2000; Kim et al. 2003). Of the 9/16 PRDM genes examined to date, all have been suggested to have growth regulatory or tumor suppressor properties (Deng and Huang 2004). In addition, some PRDM proteins appear to have roles in specification and differentiation (Morishita et al. 1992; Roy and Ng 2004). Recently, *PRDM12* has been implicated in the pathogenesis of chronic myeloid leukemia as it is found in a minimal commonly deleted region in CML patients, but its molecular function remains unknown (Kolomietz et al. 2003; Reid and Nacheva 2004)

Our analysis of Prdm12 in skeletal myogenesis provides a foundation for further functional characterization of this protein. We show that Prdm12 overexpression promotes C2C12 muscle cell differentiation in association with an elevation of d- and tMeK20-H4. Interestingly, disruption of the carboxyl terminal zinc-funger domains of Prdm12 results in its redistribution into subnuclear ring-shaped aggregates elevated in dMeK20-H4. An exciting possibility from this result is the SET domain of Prdm12 may be involved in dMeK20-H4 catalysis. One could speculate that the flag-Prdm12 aggregates result from continued redistribution to Prdm12 SET domain binding sites that undergo dMeK20-H4 catalysis. Interestingly, truncation of the carboxyl terminal zinc-finger domains of Prdm12 results in its redistribution into subnuclear ring-shaped aggregates elevated in dMeK20-H4. Whether the flag-Prdm12 engineered splice variant occurs *in vivo* for a specific function will remain an important question to address. Of note, the published subnuclear distribution of overexpressed Prdm5 appears to be similar to Prdm12 with circular aggregates distributed in DAPI-depleted regions of the nucleus (Deng and Huang 2004). It is also intriguing that the flag-Prdm12 construct prevents C2C12 differentiation and as such, it may be acting as a dominant negative mutant or have a different activity. A more detailed structure-function analysis of Prdm12 as performed above will likely provide important insights the function of Prdm12.

Overexpression of the PR-domain containing gene *Evi1* (*Prdm3*) in myeloid cells prevents their terminal differentiation into granulocytes (Morishita et al. 1992). In addition, Prdm1 (Blimp-1) is involved in cell lineage specification of neurons, and its overexpression in myoblasts can induce slow-twitch muscle differentiation (Roy and Ng 2004). Taken

together with the apparent lineage-restricted expression of several *PRDM* genes (Table 5-1), members of the PRDM family may commonly function in a cell-type specific fashion. In this regard, the repression of muscle differentiation by Prdm12 overexpression may be a reflection of its function. Given the association of flag-Prdm12 with dMeK20-H4, the normal role of Prdm12 could be to mediate gene regulation in a way that alters dMeK20-H4 of muscle-specific genes. More detailed characterization of Prdm12 could explore this possibility.

Chapter 6 ♦ Conclusions and Future Perspectives

Discussion

Histone lysine methylations demarcate distinct nuclear subdomains

The understanding of the anatomy of the nucleus is becoming increasingly complex and models of gene regulation are evolving into a three-dimensional framework (Handwerger and Gall 2006). As we dissect the various domains of the nucleus, a more physiologically relevant picture of gene regulation emerges – one that includes mobile macromolecular protein complexes in flux with a dynamically folding chromatin filament. For instance, PML domains are mobile, heterogeneous protein enrichments that make contacts with chromatin and have been suggested to coordinate various cellular processes such as DNA repair, transcription and protein degradation. In addition, splicing speckle compartments are found in a steady-state dynamic, whereby splicing proteins rapidly move in and out of these domains to various locations within the nucleoplasm. The splicing speckle periphery is enriched in transcriptional activity and, consistent with this organization, we have shown that there is a stable enrichment of tMeK4-H3 in this region (Figure s A-1 and 6-1). The proximity of active genes with splicing speckle domains could provide a more efficient means of orchestrating co-transcriptional splicing. It is remarkable that MeK36-H3 is also associated with transcriptional activity and that its corresponding HKMTase Set2 has an association with pre-mRNA splicing factors (Kizer et al. 2005). In this regard, it would be interesting to determine whether MeK36-H3 and its associated HKMTase(s) are enriched at the splicing speckle periphery.

Heterochromatic domains are also generally found to have characteristic nuclear distributions, which are preferentially associated with the nuclear and nucleolar periphery.

Consistent with the observations of others, we have found that pericentromeric heterochromatin domains are enriched in tMeK9-H3 and tMeK20-H4 (Figure 6-1). Despite the apparent overlap between these two modifications, our observations indicate that they have distinct behaviors, suggesting that they may also be spatially distinct at higher resolution. In proliferating cells of developing mid-gestation mice, tMeK9-H3 is significantly elevated in mitotic chromosomes - an association that is not observed with tMeK20-H4 (Chapter 3). Conversely, we see that over the course of differentiation in several cell types, the interphase levels of tMeK20-H4 in pericentromeric heterochromatin are progressively elevated. This association is not found with tMeK9-H3, nor is it seen with A-T rich DNA binding dyes, which co-localize with tMeK9-H3. Distinct roles for tMeK9-H3 and tMeK20-H4 are also supported by the work of others. tMeK20-H4 and tMeK9-H3 have been shown to be enriched in distinct repetitive elements (Martens et al. 2005). Moreover, although both tMeK20-H4 and tMeK9-H3 HKMTases have been shown to interact with Rb, Rb TKO (triple knockout) cells have normal tMeK9-H3 levels but depleted tMeK20-H4, suggesting these modifications behave differently in Rb pathways. It is remarkable that TKO cells have Suv39h-like anomalies such as chromosome segregation defects, which could suggest that tMeK20-H4 and tMeK9-H3 share some functions with respect to heterochromatin structure. The targeted loss of *Suv4-20h* enzymes in mice could help address more clearly any functional differences between tMeK9-H3 and tMeK20-H4. Because *Suv4-20h1* and *Suv4-20h2* have significant sequence disparity, it will be interesting to see whether these enzymes have distinct functions, which could be mediated through tMeK20-H4.

We have found that proliferating cells have localized enrichments in mMeK20-H4, which others have recently demonstrated to be facultative heterochromatin of the inactive X-chromosome (Figure 6-1). Interestingly, these enrichments are cell-cycle associated and Pr-Set7, a K20-H4 monomethylase, is important for mitosis suggesting that this enzyme may have a role in X-inactivation. In addition, Pr-Set7 functions in concert with Ezh2-dependent tMeK27-H3 on the X_i, which has also been shown to be cell-cycle regulated (Caretti et al. 2004). Consistent with such a scenario, the expression of *Pr-Set7* and *Ezh2* are similar during myogenesis (Chapter 5). An important question remains as to why there may be a wide-scale alteration in X_i- associated methylation that is cell-cycle dependent. Perhaps facultative heterochromatin plays an important role in regulating the expression of genes for differentiation. A recent study looking at the role of Eed and histone methylation in X-innactivation suggests this may be the case. Eed is part of the polycomb repressive complex 2 (PRC2), which includes Ezh2 and mediates the deposition of X-innactivating chromatin modifications including MacroH2A, tMeK27-H3 and mMeK20-H4 in trophectoderm stem cells (Kalantry et al. 2006). Kalantry et al. suggest that these histone modifications prevent differentiation by maintaining a repressive chromatin structure on the X-chromosome. Could such a mechanism be at play to regulate the mMeK20-H4 changes in muscle differentiation seen here?

Epigenetic control by tMeK20-H4

It is still unclear how long-term epigenetic information may be propagated over several cell generations but key players in this process are emerging. We have recognized a histone methyl-lysine modification, tMeK20-H4, which has several characteristics of a prototypical

regulator of epigenetic memory. Firstly, this modification is associated with gene repression, which has classically been associated with long-term memory. Importantly, the histone H4 amino terminal tail region has been proposed to mediate inter-nucleosomal interactions and tMeK20-H4 may therefore regulate stable higher-order chromatin structures. It is possible that mMeK20-H4 and dMeK20-H4 generate distinct heterochromatic structures through a similar mechanism. Secondly, tMeK20-H4 has the potential to be very stable, as an H4 variant does not appear to be present and a tMeK20-H4 specific demethylase remains elusive. Thirdly, histone H4-K20 methylations undergo considerable changes during differentiation (Figure 6-2), suggesting epigenetic changes are being encrypted in parallel with developmental progression. Fourthly, the enrichment of tMeK20-H4 over the course of differentiation is specific to certain cell types, including striated muscle, motor neurons and retinal ganglion cells. Although tMeK20-H4 levels may be enriched in other differentiating cell types, this was shown not to occur in chrondrocytic cells developing simultaneously with muscle and in other differentiated cells of the retina. Taken together, tMeK20-H4 provides important lineage-specific epigenetic memory and further efforts should be taken to elucidate the molecular factors involved in regulating this modification. In this regard, the development of Suv4-20h1/2 deficient mice should provide mechanistic insight.

Smyd1 and Prdm12 in myogenesis

In the fifth chapter of this thesis, we extend our analysis of histone lysine methylation in myogenesis to investigate its mechanistic basis. We hypothesized that muscle specific HKMTases could underlie the histone methylation changes we identified during muscle

differentiation, and that these genes would be restricted in this tissue. By *in silico* expression analysis, we identified *Prdm12* and *Smyd1* as potential HKMTases for myogenesis. Our localization data have provided some insight into their possible nuclear functions, especially in the case of Prdm12, which will be important for further molecular dissestion of these proteins.

Interestingly, Smyd1 overexpression appeared to repress skeletal muscle differentiation in embryonic muscle cultures but not in C2C12 myoblasts. This result may reflect the different levels of temporal commitment in the muscle lineage of our primary cultures versus C2C12 cells. We previously observed that primary muscle cultures derived from E10.5 embryos had detectable levels of Pax3, an early regulator of myogenesis, but Pax3 was no longer detectable following ~ 3 days of culture (data not shown). In contrast, Pax3 is not detectable in either proliferating or differentiating C2C12s, suggesting this cell line is in a later stage of muscle development than our embryonic muscle cultures. In this regard, the repressive effect of Smyd1 overexpression in primary cultures could suggest a specific temporal requirement for its function in skeletal muscle differentiation. It is possible that early overexpression of *Smyd1* causes ectopic activation or repression of cell-cycle or differentiation genes, respectively.

The subnuclear distribution of Smyd1 and apparent role in myogenesis is important in light of chromatin dynamics occurring in active gene compartments. Moen et al. demonstrated a striking repositioning of some myogenic genes to active chromatin compartments at the periphery of splicing speckles that occurred specifically during differentiation of skeletal muscle (Moen et al. 2004). Importantly, we have observed that

the periphery of splicing speckles is significantly enriched in tMeK4-H3. Therefore it is tempting to speculate that Smyd1 may function to dynamically regulate myogenic genes in tMeK4-H3-rich chromatin domains.

Our experiments with Prdm12 have provided interesting possibilities of its function that would mandate further characterization in order to gain a solid understanding of its biological role. It would be important to substantiate the dMeK20-H4 specificity of Prdm12 by *in vitro* methyltransferase assays, as it has been done for other HKMTases. In addition, it would be informative to investigate the endogenous expression profile of *Prdm12* during embryogenesis *in situ* and to perform mouse, xenopus, or zebrafish knockout studies for further characterization of this HKMTase.

An epigenetic blueprint of mammalian embryogenesis

In recent years, reports showing the unique distributions of histone methyl-lysine derivatives during development have surfaced. These studies demontrate that histone lysine methylation is dynamic during development and supports the idea that these modifications are key epigenetic regulators. Importantly, changes involving specific histone methyl-lysines occur on a broad scale readily identifiable by immunofluorescence. This has been shown during mouse pre-implantation development, germ cell specification and through our data during mid-gestation. Although the distributions of various methylations remain to be examined in several developmental events, an integrated understanding of the epigenome during mammalian development could evolve from these studies.

Mouse pre-implantation development

During pre-implantation development, epigenetic reprogramming events occur that are thought to facilitate the expression of embryonic genes (Forneris et al. 2005). From the early stages of fertilization, maternal metaphase II chromatin is enriched in d/tMeK9-H3, whereas these modifications are absent or extremely depleted in decondensing sperm nuclei (Sarmento et al. 2004). These strikingly different methylation profiles are also apparent in male and female pronuclei, and persist until the 2-cell stage (Liu et al. 2004). Consistently, tMeK20-H4 has also been shown to be present in female pronuclei but absent in paternal pronuclei (Kourmouli et al. 2004). *De novo* methylation of histone H3-K9 occurs at the 4-cell stage, causing the asymmetry between male and female genomes to be lost (Liu et al. 2004). This change is thought to occur by alleviation of methylase repression from the paternal genome. In addition, a global depletion in methylation of R17-H3, R3-H4 and acH4 is found on metaphase II egg chromatin and persists in early embryonic blastomeres (Sarmento et al. 2004). Interestingly, MeR17-H3 depletion occurs through a peptidyl arginine deiminase (PAD) thought to be present in the maternal cytoplasm. At the blastocyst stage, detectable levels of MeR3-H4 and AcH4 are seen in metaphase blastomeres. Together, these studies suggest significant changes in histone methylation are important for the earliest steps in mammalian development.

Dynamic changes in histone methylation during early cleavage stages have recently been shown to control imprinted inactivation of the paternal X-chromosome (Xp). By the time K9-H3 levels have equilibrated between maternal and paternal genomes at the 4-cell stage, Xist RNA begins to coat the Xp (Okamoto et al. 2004). Importantly, the timing of

this event enables a distinction between both genomes to be maintained. At the 8-cell stage, the Xp is hypomethylated at K4-H3 and hypoacetylated at K9-H3. This is followed by deposition of d/tMeK27-H3 through EED/Enx1 accumulation by the morula stage (16-cell) (Okamoto et al. 2004). At the blastocyst stage, dMeK9-H3 is finally deposited on the Xp. However, a striking disassembly of Xp inactivating components occurs during inner cell mass growth and, by implantation, the random inactivation of paternal and maternal X chromosomes begins. These dynamic changes reveal a remarkable plasticity of histone lysine methylation marks during pre-implantation development.

Epigenetic reprogramming in mouse germ cell development

Mouse germ cells are derived from epiblast cells, which migrate into the gonads during mid-gestation to undergo extensive epigenetic reprogramming. Recently, histone methyl-lysine modifications have been shown to undergo considerable dynamics in developing germ cells (Seki et al. 2005). During germ cell specification at E8, a rapid and dramatic loss of dMeK9-H3 is seen in germ cells, whereas tMeK9-H3 remains relatively stable (Seki et al. 2005). From E8.5 onward, there appears to be a progressive increase in tMeK27-H3 in euchromatin (Seki et al. 2005). Interestingly, tMeK27-H3 has also been shown to accumulate on the X_i during development and appears to generate a repressive but plastic chromatin state (Plath et al. 2003; Silva et al. 2003). In addition, MeK4-H3 and acK9-H3 remain relatively stable throughout germ cell specification but are globally elevated as germ cells populate the genital ridges. Although enrichments in these modifications generate a more euchromatic state there is no elevation in genome-wide transcription, suggesting these germ cells simply have decondensed chromatin (Seki et al. 2005). These

epigenetic resetting events are thought to be necessary to generate a totipotent gamete with less somatic character. These recent findings in germ cell development could provide a model to enable the understanding of histone lysine demethylation of repressive chromatin marks.

Mouse mid-gestation development

We have provided the first report showing genome-wide changes in histone methyl-lysine derivatives in mammalian embryos from E8.5 to E11.5. Our work therefore supplements the studies above in providing a description of histone methylation changes during mid-gestation. In proliferating cells during this timescale, mMeK20-H4 was elevated in both mitotic chromosomes and interphase chromatin. Similarly, tMeK9-H3 was elevated in mitotic chromosomes, demonstrating that these methyl-lysines are associated with cell cycle events during embryogenesis. The most striking finding was that tMeK20-H4 levels are progressively elevated in post-mitotic cells of differentiating motor neurons, cardiac muscle and skeletal muscle cells, beginning at E9.5. Based on the association of tMeK27-H3 with the Xi, we would expect this modification to parallel the distributions of mMeK20-H4 during embryogenesis. It would be important to examine other modifications that have not been examined in our study, and to look at earlier and later events of embryogenesis in order to get a more complete picture of epigenetic events in development. Nevertheless, our study provides a foundation for examining the role of histone lysine methylation during mid-gestation that could be used to examine the chromatin of HKMTase-deficient mice.

Future perspectives

HKMTases and diet

The involvement of dietary methyl donors in cancer has been known for quite some time, however the molecular causes underlying this link have been largely elusive (Huang 2002). The finding that HKMTases transfer methyl groups obtained from S-adenosylmethionine (SAM) derived from methyl donors such as choline, methionine, and folate, provides an epigenetic link to the nutritional causes of cancer.

A decrease in methyl donors can result in decreased SAM and elevated S-adenosylhomocysteine (SAH), which can impede the methionine cycle (Huang 2002). This can be counteracted by dietary intake of methyl donors, vitamins B12 and B6 (Huang 2002). However, some individuals may have hypomorphic alleles of key enzymes in the methionine cycle, such as methionine synthase (MS) or methylenetetrahydrofolate reductase (MTHFR), thereby predisposing them to cancer (Huang 2002). In addition, having a genetic deficiency in the methionine pathway and a methyl donor depleted diet compounds the risk of cancer (Huang 2002). In this regard, individuals with HKMTase deficiency such as Sotos syndrome could have a decreased risk of developing cancer through dietary supplementation.

Epigenetics of retinal development

In chapter 3, we provided a foundation for exploring histone lysine methylation-dependent epigenetic changes in retinogenesis. To our knowledge, this is the first report describing broad histone methylation changes during the development of the vertebrate retina. We

demonstrated that tMeK20-H4 is progressively enriched in cells of the differentiating ganglion cell layer, but not other differentiated retinal cells, providing support for a cell-type specific role of tMeK20-H4 during differentiation. We postulate that the loss of tMeK20-H4 in non retinal ganglion cells of the retina could be due to an unidentified demethylase specific to these cells or absence of a HKMTase, resulting in mitotic dilution. We predict that a retinal-specific HKMTase could account for ganglion cell-specific elevation of tMeK20-H4. *PRDM8* appears to be an excellent candidate in this regard, as it is expressed in the retina by *in situ*, consistent with *in silico* expression data showing that this HKMTase is elevated in the eye (Chapter 5). It is noteworthy that PRDM proteins are generally tissue-restricted, which could implicate these genes in the regulation of tissue-specific histone lysine methylation.

Our data may provide important insight towards an understanding of epigenetic plasticity in retinal stem cells. It would be interesting to determine the role of tMeK20-H4 and other histone methylations in this context. Based on our findings, it would also be interesting to determine whether the heterogeneity in the retinal precursor pool can be explained in epigenetic terms. Histone lysine methylation could be examined with respect to the spatiotemporal occurrence of bHLH transcription factors, such as NeuroD and Math5, regulators that have been shown to influence lineage restriction of RPCs.

Histone methylation in epigenetic disorders

The role of histone lysine methylation in epigenetics has been largely elusive until recent years and, perhaps for this reason, it has not been thoroughly examined in human epigenetic diseases. In contrast, a more detailed understanding of DNA methylation has

lead to the identification of several epigenetic diseases that are caused by deregulation of this process. Given the interrelationships between epigenetic modifications, histone methylation may play an important role in disorders that have disrupted DNA methylation. For instance, a recent analysis of cancer cells with lower than normal levels of 5-methylcytosine also identified a dramatic reduction of acK16-H4 and tMeK20-H4 in distinct repetitive elements, namely NBL2, Satellite 2 and D4Z4 sequences (Fraga et al. 2005). In this light, it is interesting that disruption of D4Z4 repeats is associated with fascioscapulohumeral muscular dystrophy (FSHD) (Buzhov et al. 2005a; Buzhov et al. 2005b). In addition, Satellite II sequences associated with pericentromeric heterochromatin are DNA hypomethylated in immunodeficiency, centromeric instability, facial anomalies syndrome (ICF), which is often caused by mutations in the DNA methyltransferase DNMT3b (Gisselsson et al. 2005). The pathological association of tMeK20-H4 with D4Z4 and Satellite 2 repeats may implicate this modification in the etiology of FSHD or ICF. Histone lysine methylation may also have a role in imprinting disorders, as it has been demonstrated for some cases of Beckwith-Weideman syndrome caused by the HKMTase NSD1 (Baujat et al. 2004). Data presented in this thesis would suggest that defects in tMeK20-H4 could be implicated in the molecular basis of disorders affecting the nervous system, cardiac and skeletal muscle, and vision. A role for tMeK20-H4 has already been shown to be important in human cancers (Fraga et al. 2005). In this regard, cellular histone methylation profiles may also serve to predict the clinical outcome of epigenetic disorders as it has been recently shown with prostate cancer (Seligson et al. 2005).

Histone methyl-lysines as chemotherapeutic targets

Given the pervasive role of HKMTases in carcinogenesis, histone lysine methylation is an attractive chemotherapeutic target. Other epigenetic modifications such as DNA methylation and histone acetylation have proven to be successful drug targets for certain types of human cancers. For instance DNA methylation inhibitors such as 5'azacytidine are being used in phase I-III clinical trials for leukemias while more efficacious analogues of this drug are being developed (Egger et al. 2004; Bhalla 2005). Similarly, a variety of HDAC inhibitors are being successfully used to treat some forms of cancer in clinical trials (Egger et al. 2004; Espino et al. 2005). The influence of distinct histone lysine methylations on specific chromatin domains could make this a more promising epigenetic target for cancer drugs. Because histone methylation marks are thought to be relatively stable, it is conceivable that drugs capable of modifiying active or heterochromatic methylations could be very efficient, with less need for repetitive treatments as is one downfall of HDAC inhibitors. In addition, histone methylating/demethylating agents could be used in conjunction with current epigenetic therapies to enhance their efficacy.

Concluding remarks

This thesis has added a small piece to an epigenetic puzzle that becomes increasingly complex with growing knowledge of histone post-translational modifications and DNA methylation. Epigenetic phenomena hold great promise for gaining a better functional understanding of the genome, which could materialize into new treatments for human diseases. Large scale analyses such as the human epigenome project are likely to propel new and exciting epigenetic research in years to come.

Figure 6- 1: Functional chromatin domains defined by histone lysine methylation. Pericentromeric heterochromatin in mammalian nuclei is enriched in tMeK9-H3 and tMeK20-H4. mMeK20-H4 is enriched in facultative heterochromatin, defining transcriptionally repressive chromatin environments. The periphery of spicling speckle compartments is enriched in tMeK4-H3, a region also shown to be elevated in hyperacetylated H3, defining a transcriptionally active subnuclear environment. Pericentromeric heterochromatin compartments increase in size during differentiation of particular cell types, with a corresponding elevation of tMeK20-H4. In contrast, elevated levels of mMeK20-H4 are associated with proliferation.

H3 trimethyl K4
-transcriptional
activation
H4 trimethyl K20
-transcriptional
repression
-cell differentiation
H4 monoMeK20
-cell proliferation
-facultative heterochromatinization
H3 trimethyl K9
-transcriptional
repression
-cell differentiation ?

Figure 6- 2: Subnuclear re-organization of chromatin through dynamic K20-H4 methylation during muscle differentiation.

Skeletal muscle differentiation of C2C12 cells is marked by changes in the subnuclear distribution of K20-H4 methylations. In proliferating mononucleated cells, ~95 % of nuclei have 2 large subnuclear aggregates of mMeK20-H4 (red) which abut either nucleoli or the nuclear periphery, nucleolar enrichment of dMeK20-H4 (green) and a pericentric heterochromatin domain (blue) enriched in tMeK20-H4. Induction of myogenic differentiation by serum withdrawal causes a significant depletion of large mMeK20-H4 aggregates and loss of dMeK20-H4 nucleolar enrichment, while retaining other punctate foci, with concomitant enlargement of tMeK20-H4 pericentric heterochromatin domains. In multi-nucleated, terminally differentiated cells, tMeK20-H4 domains are significantly enlarged, while only punctate foci remain for mMeK20-H4 and dMeK20-H4. These dynamic changes in K20 methylation represent major epigenetic changes that are suggestive of chromatin re-organization over the course of differentiation.

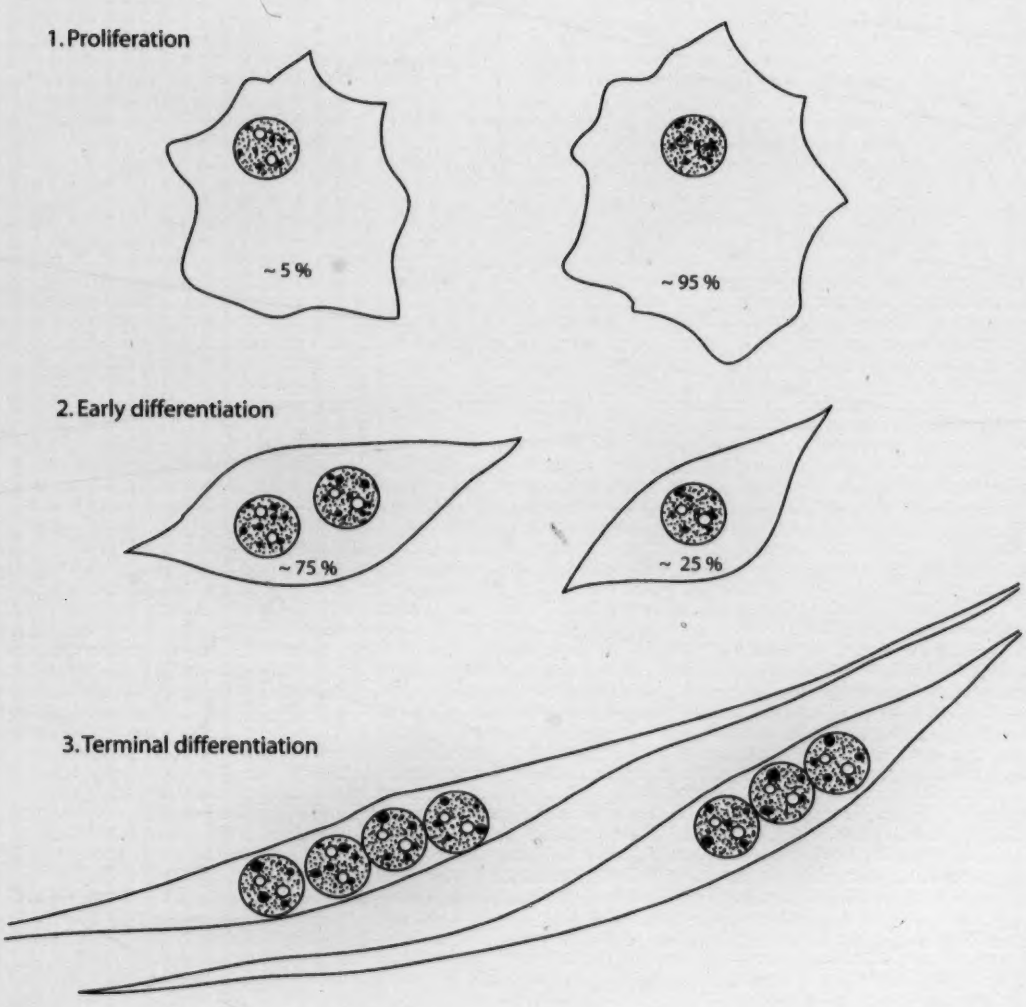

1. Proliferation
~ 5 %
~ 95 %
2. Early differentiation
~ 75 %
~ 25 %
3. Terminal differentiation

Table 6- 1: Summary of histone lysine methylation changes occurring during mouse embryogenesis.

Developmental stage	Histone methyl-lysine change	References
E 1.0 (Fertilization)	-Maternal genome: elevated d/tMeK9-H3 and tMeK20-H4, decreased H4ac -Paternal genome: elevated H4ac, decreased d/tMeK9-H3 and tMeK20-H4	(Kourmouli et al. 2004; Liu et al. 2004; Sarmento et al. 2004)
E 1.5 (2-cell)	-Maternal genome: elevated d/tMeK9-H3, decreased H4ac -Paternal genome: elevated acH4, decreased d/tMeK9-H3 -Global decrease in meR17-H3, meR3-H4 and acH4	(Liu et al. 2004; Sarmento et al. 2004)
E 2.0 (4-cell)	-d/tMeK9-H3 asymmetry lost -Global decrease in meR17-H3, meR3-H4 and acH4	(Liu et al. 2004; Sarmento et al. 2004)
E 2.5-3.0 (8-cell)	-Global decrease in meR17-H3, meR3-H4 and acH4 Xp: decreased MeK4-H3 and acK9-H3 followed by elevated K27-H3me	(Okamoto et al. 2004; Sarmento et al. 2004)
E 4.0 (Early blastocyst)	Appearance of meR3H3 and acH4 -Xp: rise in dMeK9-H3 levels	(Okamoto et al. 2004; Sarmento et al. 2004)
E 4.5 (Implantation)	-Dissasembly of Xp inactivation during inner cell mass growth and random X-inactivation	(Okamoto et al. 2004)
E 8.5 (Gastrulation)	-Proliferating cells elevated in tMeK9-H3 and mMeK20-H4 -Loss of dMeK9-H3 and elevation of tMeK27-H3 in primordial germ cells	(Biron et al. 2004; Seki et al. 2005)
E 9.5 (Gastrulation)	-Proliferating cells elevated in tMeK9-H3 and mMeK20-H4 -Progressive elevation of tMeK20-H4 in differentiating motor neurons and striated muscle cells -Progressive elevation of tMeK27-H3 in migrating germ cells	(Biron et al. 2004; Seki et al. 2005)
E 10.5 (Gastrulation)	-Proliferating cells elevated in tMeK9-H3 and mMeK20-H4 -Overall elevation of tMeK20-H4 and depletion of mMeK20-H4 in differentiated striated muscle, ventrolateral motor neurons and retinal ganglion cells (chick) -Global elevation of MeK4-H3 and acK9-H3 in germ cells of the genital ridge	(Biron et al. 2004; Seki et al. 2005)

Appendices

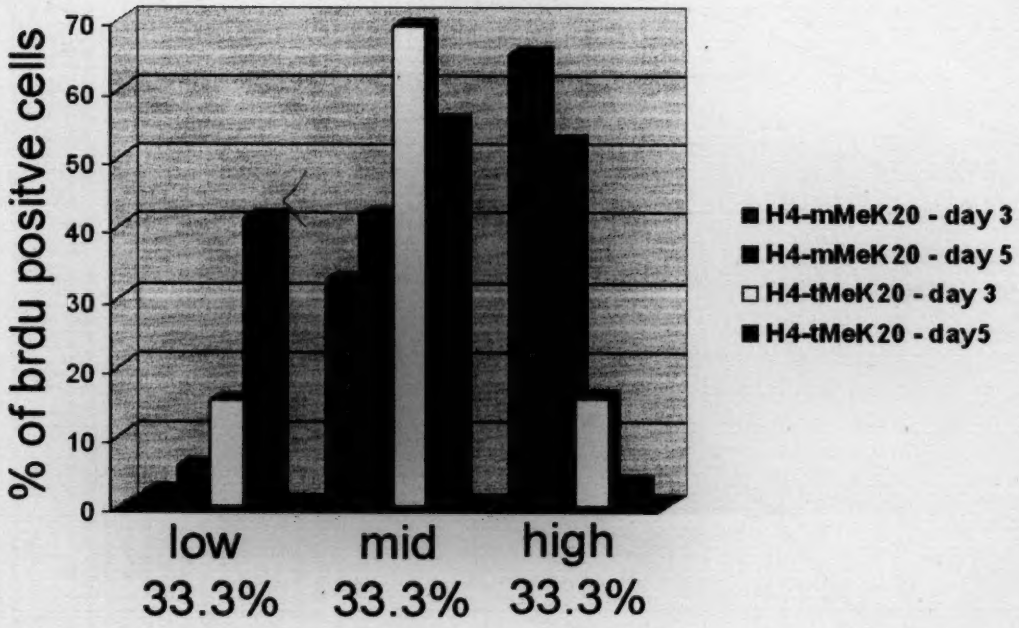
Figure A- 1: Subnuclear distribution of tMeK4-H3 in COS-7 cells and mouse embryo sections.

(A) Co-immunofluorescence of SC-35 (green) and tMeK4-H3 (red) in COS-7 nuclei. The relative distribution of SC-35 and tMeK4-H3 is shown by linescan through the merged image. The tMeK4-H3 threshold panel displays the extensive association of higher intensity tMeK4 values (top 66%) with SC-35. (B) Subnuclear distribution of tMeK4-H3 (red) relative to DAPI (green) in mouse E10.5 sections. Individual channels are shown in grayscale and the merged image is pseudocolored.

A
SC-35
tMeK4
merge
0
250
B
DAPI
tMeK4
merge
tMeK4 threshold
Position

Figure A- 2: Cell cycle analysis of K20-H4 methylations in differentiating mouse primary limb bud cultures.

Mouse primary limb bud cultures derived from E10.5 embryos were pulsed with BrdU for 2 hours following 3 or 5 days of culture and processed for co-immunofluorescence with BrdU and mono- or tri-methyl K20 H4. Greater than 50 BrdU positive nuclei were scored from 40x images for each co-immunofluorescence stain and assigned in either the lower, middle, or upper third of intensity values.


% of brdu positve cells
70
60
50
40
30
20
10
0
low
33.3%
mid
33.3%
high
33.3%
H4-mMeK20 - day 3
H4-mMeK20 - day 5
H4-tMeK20 - day 3
H4-tMeK20 - day5

References

Ait-Si-Ali S, Guasconi V, Fritsch L, Yahi H, Sekhri R, Naguibneva I, Robin P, Cabon F, Polesskaya A, Harel-Bellan A (2004) A Suv39h-dependent mechanism for silencing S-phase genes in differentiating but not in cycling cells. Embo J 23:605-615

Al-Mulla N, Belgaumi AF, Teebi A (2004) Cancer in Sotos syndrome: report of a patient with acute myelocytic leukemia and review of the literature. J Pediatr Hematol Oncol 26:204-208

Allis CD, Bowen JK, Abraham GN, Glover CV, Gorovsky MA (1980) Proteolytic processing of histone H3 in chromatin: a physiologically regulated event in Tetrahymena micronuclei. Cell 20:55-64

Angrand PO, Apiou F, Stewart AF, Dutrillaux B, Losson R, Chambon P (2001) NSD3, a new SET domain-containing gene, maps to 8p12 and is amplified in human breast cancer cell lines. Genomics 74:79-88

Arber S, Han B, Mendelsohn M, Smith M, Jessell TM, Sockanathan S (1999) Requirement for the homeobox gene Hb9 in the consolidation of motor neuron identity. Neuron 23:659-674

Arney KL, Bao S, Bannister AJ, Kouzarides T, Surani MA (2002) Histone methylation defines epigenetic asymmetry in the mouse zygote. Int J Dev Biol 46:317-320

Ayyanathan K, Lechner MS, Bell P, Maul GG, Schultz DC, Yamada Y, Tanaka K, Torigoe K, Rauscher FJ, 3rd (2003) Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation. Genes Dev 17:1855-1869

Bannister AJ, Schneider R, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T (2005) Spatial distribution of Di- and tri-methyl lysine 36 of histone H3 at active genes. J Biol Chem 280:17732-17736

Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, Kouzarides T (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature 410:120-124

Baujat G, Rio M, Rossignol S, Sanlaville D, Lyonnet S, Le Merrer M, Munnich A, Gicquel C, Cormier-Daire V, Colleaux L (2004) Paradoxical NSD1 mutations in Beckwith-Wiedemann syndrome and 11p15 anomalies in Sotos syndrome. Am J Hum Genet 74:715-720

Baujat G, Rio M, Rossignol S, Sanlaville D, Lyonnet S, Merrer ML, Munnich A, Gicquel C, Colleaux L, Cormier-Daire V (2005) Clinical and molecular overlap in overgrowth syndromes. Am J Med Genet C Semin Med Genet

Bayne EH, Allshire RC (2005) RNA-directed transcriptional gene silencing in mammals. Trends Genet 21:370-373

Beisel C, Imhof A, Greene J, Kremmer E, Sauer F (2002) Histone methylation by the Drosophila epigenetic transcriptional regulator Ash1. Nature 419:857-862

Belmont A (2003) Dynamics of chromatin, proteins, and bodies within the cell nucleus. Curr Opin Cell Biol 15:304-310

Belmont AS (2001) Visualizing chromosome dynamics with GFP. Trends Cell Biol 11:250-257

Bergemann AD, Cole F, Hirschhorn K (2005) The etiology of Wolf-Hirschhorn syndrome. Trends Genet 21:188-195

Berger SL (2002) Histone modifications in transcriptional regulation. Curr Opin Genet Dev 12:142-148

Berkes CA, Tapscott SJ (2005) MyoD and the transcriptional control of myogenesis. Semin Cell Dev Biol 16:585-595

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125:315-326

Bhalla KN (2005) Epigenetic and chromatin modifiers as targeted therapy of hematologic malignancies. J Clin Oncol 23:3971-3993

Bird G, Zorio DA, Bentley DL (2004) RNA polymerase II carboxy-terminal domain phosphorylation is required for cotranscriptional pre-mRNA splicing and 3'-end formation. Mol Cell Biol 24:8963-8969

Biron VL, McManus KJ, Hu N, Hendzel MJ, Underhill DA (2004) Distinct dynamics and distribution of histone methyl-lysine derivatives in mouse development. Dev Biol 276:337-351

Blais A, Tsikitis M, Acosta-Alvear D, Sharan R, Kluger Y, Dynlacht BD (2005) An initial blueprint for myogenic differentiation. Genes Dev 19:553-569

Borlak J, Thum T (2003) Hallmarks of ion channel gene expression in end-stage heart failure. Faseb J 17:1592-1608

Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK, Bell GW, Otte AP, Vidal M, Gifford DK, Young RA, Jaenisch R (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature 441:349-353

Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K (2006) Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev 20:1123-1136

Brehm A, Tufteland KR, Aasland R, Becker PB (2004) The many colours of chromodomains. Bioessays 26:133-140

Brero A, Easwaran HP, Nowak D, Grunewald I, Cremer T, Leonhardt H, Cardoso MC (2005) Methyl CpG-binding proteins induce large-scale chromatin reorganization during terminal differentiation. J Cell Biol 169:733-743

Brown KE, Amoils S, Horn JM, Buckle VJ, Higgs DR, Merkenschlager M, Fisher AG (2001) Expression of alpha- and beta-globin genes occurs within different nuclear domains in haemopoietic cells. Nat Cell Biol 3:602-606

Brown MA, Sims RJ, 3rd, Gottlieb PD, Tucker PW (2006) Identification and characterization of Smyd2: a split SET/MYND domain-containing histone H3 lysine 36-specific methyltransferase that interacts with the Sin3 histone deacetylase complex. Mol Cancer 5:26

Buzhov BT, Lemmers RJ, Tournev I, Dikova C, Kremensky I, Petrova J, Frants RR, van der Maarel SM (2005a) Genetic confirmation of facioscapulohumeral muscular dystrophy in a case with complex D4Z4 rearrangments. Hum Genet 116:262-266

Buzhov BT, Lemmers RJ, Tournev I, van der Wielen MJ, Ishpekova B, Petkov R, Petrova J, Frants RR, Padberg GW, van der Maarel SM (2005b) Recurrent somatic mosaicism for D4Z4 contractions in a family with facioscapulohumeral muscular dystrophy. Neuromuscul Disord 15:471-475

Canote R, Du Y, Carling T, Tian F, Peng Z, Huang S (2002) The tumor suppressor gene RIZ in cancer gene therapy (review). Oncol Rep 9:57-60

Cao R, Zhang Y (2004) The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. Curr Opin Genet Dev 14:155-164

Caretti G, Di Padova M, Micales B, Lyons GE, Sartorelli V (2004) The Polycomb Ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation. Genes Dev 18:2627-2638

Carmo-Fonseca M (2002) The contribution of nuclear compartmentalization to gene regulation. Cell 108:513-521

Cash DE, Bock CB, Schughart K, Linney E, Underhill TM (1997) Retinoic acid receptor alpha function in vertebrate limb skeletogenesis: a modulator of chondrogenesis. J Cell Biol 136:445-457

Cerveira N, Correia C, Doria S, Bizarro S, Rocha P, Gomes P, Torres L, Norton L, Borges BS, Castedo S, Teixeira MR (2003) Frequency of NUP98-NSD1 fusion transcript in childhood acute myeloid leukaemia. Leukemia 17:2244-2247

Chaly N, Munro SB (1996) Centromeres reposition to the nuclear periphery during L6E9 myogenesis in vitro. Exp Cell Res 223:274-278

Charier G, Couprie J, Alpha-Bazin B, Meyer V, Quemeneur E, Guerois R, Callebaut I, Gilquin B, Zinn-Justin S (2004) The Tudor tandem of 53BP1: a new structural motif involved in DNA and RG-rich peptide binding. Structure (Camb) 12:1551-1562

Cheng X, Collins RE, Zhang X (2005) Structural and sequence motifs of protein (histone) methylation enzymes. Annu Rev Biophys Biomol Struct 34:267-294

Ching RW, Dellaire G, Eskiw CH, Bazett-Jones DP (2005) PML bodies: a meeting place for genomic loci? J Cell Sci 118:847-854

Chuikov S, Kurash JK, Wilson JR, Xiao B, Justin N, Ivanov GS, McKinney K, Tempst P, Prives C, Gamblin SJ, Barlev NA, Reinberg D (2004) Regulation of p53 activity through lysine methylation. Nature 432:353-360

Cocklin RR, Wang M (2003) Identification of methylation and acetylation sites on mouse histone H3 using matrix-assisted laser desorption/ionization time-of-flight and nanoelectrospray ionization tandem mass spectrometry. J Protein Chem 22:327-334

Collins KA, Furuyama S, Biggins S (2004) Proteolysis contributes to the exclusive centromere localization of the yeast Cse4/CENP-A histone H3 variant. Curr Biol 14:1968-1972

Cosgrove MS, Wolberger C (2005) How does the histone code work? Biochem Cell Biol 83:468-476

Craig JM (2005) Heterochromatin--many flavours, common themes. Bioessays 27:17-28

Cremer T, Cremer C (2001) Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet 2:292-301

Cuthbert GL, Daujat S, Snowden AW, Erdjument-Bromage H, Hagiwara T, Yamada M, Schneider R, Gregory PD, Tempst P, Bannister AJ, Kouzarides T (2004) Histone deimination antagonizes arginine methylation. Cell 118:545-553

Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V (2002) Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. Cell 111:185-196

Daser A, Rabbitts TH (2004) Extending the repertoire of the mixed-lineage leukemia gene MLL in leukemogenesis. Genes Dev 18:965-974

de la Cruz X, Lois S, Sanchez-Molina S, Martinez-Balbas MA (2005) Do protein motifs read the histone code? Bioessays 27:164-175

Deardorff MA, Maisenbacher M, Zackai EH (2004) Ganglioglioma in a Sotos syndrome patient with an NSD1 deletion. Am J Med Genet A 130:393-394

Deng Q, Huang S (2004) PRDM5 is silenced in human cancers and has growth suppressive activities. Oncogene 23:4903-4910

Denman RB (2005) PAD: the smoking gun behind arginine methylation signaling? Bioessays 27:242-246

Derunes C, Briknarova K, Geng L, Li S, Gessner CR, Hewitt K, Wu S, Huang S, Woods VI, Jr., Ely KR (2005) Characterization of the PR domain of RIZ1 histone methyltransferase. Biochem Biophys Res Commun 333:925-934

Dillon N (2004) Heterochromatin structure and function. Biol Cell 96:631-637

Dion MF, Altschuler SJ, Wu LF, Rando OJ (2005) Genomic characterization reveals a simple histone H4 acetylation code. Proc Natl Acad Sci U S A 102:5501-5506

Dodge JE, Kang YK, Beppu H, Lei H, Li E (2004) Histone H3-K9 methyltransferase ESET is essential for early development. Mol Cell Biol 24:2478-2486

Douglas J, Coleman K, Tatton-Brown K, Hughes HE, Temple IK, Cole TR, Rahman N (2005) Evaluation of NSD2 and NSD3 in overgrowth syndromes. Eur J Hum Genet 13:150-153

Ebert A, Schotta G, Lein S, Kubicek S, Krauss V, Jenuwein T, Reuter G (2004) Su(var) genes regulate the balance between euchromatin and heterochromatin in Drosophila. Genes Dev 18:2973-2983

Egger G, Liang G, Aparicio A, Jones PA (2004) Epigenetics in human disease and prospects for epigenetic therapy. Nature 429:457-463

Elgin SC, Grewal SI (2003) Heterochromatin: silence is golden. Curr Biol 13:R895-898

Erhardt S, Su IH, Schneider R, Barton S, Bannister AJ, Perez-Burgos L, Jenuwein T, Kouzarides T, Tarakhovsky A, Surani MA (2003) Consequences of the depletion of

zygotic and embryonic enhancer of zeste 2 during preimplantation mouse development. Development 130:4235-4248

Ernst P, Fisher JK, Avery W, Wade S, Foy D, Korsmeyer SJ (2004a) Definitive hematopoiesis requires the mixed-lineage leukemia gene. Dev Cell 6:437-443

Ernst P, Mabon M, Davidson AJ, Zon LI, Korsmeyer SJ (2004b) An Mll-dependent Hox program drives hematopoietic progenitor expansion. Curr Biol 14:2063-2069

Eskiw CH, Dellaire G, Bazett-Jones DP (2004) Chromatin contributes to structural integrity of promyelocytic leukemia bodies through a SUMO-1-independent mechanism. J Biol Chem 279:9577-9585

Espino PS, Drobic B, Dunn KL, Davie JR (2005) Histone modifications as a platform for cancer therapy. J Cell Biochem 94:1088-1102

Esteller M (2005) DNA methylation and cancer therapy: new developments and expectations. Curr Opin Oncol 17:55-60

Fang J, Feng Q, Ketel CS, Wang H, Cao R, Xia L, Erdjument-Bromage H, Tempst P, Simon JA, Zhang Y (2002) Purification and functional characterization of SET8, a nucleosomal histone H4-lysine 20-specific methyltransferase. Curr Biol 12:1086-1099

Faravelli F (2005) NSD1 mutations in Sotos syndrome. Am J Med Genet C Semin Med Genet

Farris SD, Rubio ED, Moon JJ, Gombert WM, Nelson BH, Krumm A (2005) Transcription-induced chromatin remodeling at the C-Myc gene involves a local exchange of histone H2A.Z. J Biol Chem

Faust C, Lawson KA, Schork NJ, Thiel B, Magnuson T (1998) The Polycomb-group gene eed is required for normal morphogenetic movements during gastrulation in the mouse embryo. Development 125:4495-4506

Feng Q, Wang H, Ng HH, Erdjument-Bromage H, Tempst P, Struhl K, Zhang Y (2002) Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. Curr Biol 12:1052-1058

Fidanza V, Melotti P, Yano T, Nakamura T, Bradley A, Canaani E, Calabretta B, Croce CM (1996) Double knockout of the ALL-1 gene blocks hematopoietic differentiation in vitro. Cancer Res 56:1179-1183

Filipowicz W, Jaskiewicz L, Kolb FA, Pillai RS (2005) Post-transcriptional gene silencing by siRNAs and miRNAs. Curr Opin Struct Biol 15:331-341

Fischle W, Wang Y, Allis CD (2003) Histone and chromatin cross-talk. Curr Opin Cell Biol 15:172-183

Fisher AG, Merkenschlager M (2002) Gene silencing, cell fate and nuclear organisation. Curr Opin Genet Dev 12:193-197

Fodor BD, Kubicek S, Yonezawa M, O'Sullivan RJ, Sengupta R, Perez-Burgos L, Opravil S, Mechtler K, Schotta G, Jenuwein T (2006) Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells. Genes Dev 20:1557-1562

Forneris F, Binda C, Vanoni MA, Battaglioli E, Mattevi A (2005) Human histone demethylase LSD1 reads the histone code. J Biol Chem 280:41360-41365

Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, Bonaldi T, Haydon C, Ropero S, Petrie K, Iyer NG, Perez-Rosado A, Calvo E, Lopez JA, Cano A, Calasanz MJ, Colomer D, Piris MA, Ahn N, Imhof A, Caldas C, Jenuwein T, Esteller M (2005) Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nat Genet 37:391-400

Francastel C, Walters MC, Groudine M, Martin DI (1999) A functional enhancer suppresses silencing of a transgene and prevents its localization close to centrometric heterochromatin. Cell 99:259-269

Freitag M, Selker EU (2005) Controlling DNA methylation: many roads to one modification. Curr Opin Genet Dev 15:191-199

Frigola J, Sole X, Paz MF, Moreno V, Esteller M, Capella G, Peinado MA (2005) Differential DNA hypermethylation and hypomethylation signatures in colorectal cancer. Hum Mol Genet 14:319-326

Garcia-Cao M, O'Sullivan R, Peters AH, Jenuwein T, Blasco MA (2004) Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. Nat Genet 36:94-99

Gisselsson D, Shao C, Tuck-Muller CM, Sogorovic S, Palsson E, Smeets D, Ehrlich M (2005) Interphase chromosomal abnormalities and mitotic missegregation of hypomethylated sequences in ICF syndrome cells. Chromosoma 114:118-126

Gonzalo S, Blasco MA (2005) Role of Rb family in the epigenetic definition of chromatin. Cell Cycle 4:752-755

Gonzalo S, Garcia-Cao M, Fraga MF, Schotta G, Peters AH, Cotter SE, Eguia R, Dean DC, Esteller M, Jenuwein T, Blasco MA (2005) Role of the RB1 family in stabilizing histone methylation at constitutive heterochromatin. Nat Cell Biol 7:420-428

Gorisch SM, Wachsmuth M, Ittrich C, Bacher CP, Rippe K, Lichter P (2004) Nuclear body movement is determined by chromatin accessibility and dynamics. Proc Natl Acad Sci U S A 101:13221-13226

Gottlieb PD, Pierce SA, Sims RJ, Yamagishi H, Weihe EK, Harriss JV, Maika SD, Kuziel WA, King HL, Olson EN, Nakagawa O, Srivastava D (2002) Bop encodes a muscle-restricted protein containing MYND and SET domains and is essential for cardiac differentiation and morphogenesis. Nat Genet 31:25-32

Grewal SI, Elgin SC (2002) Heterochromatin: new possibilities for the inheritance of structure. Curr Opin Genet Dev 12:178-187

Grishok A (2005) RNAi mechanisms in Caenorhabditis elegans. FEBS Lett 579:5932-5939

Hamamoto R, Furukawa Y, Morita M, Iimura Y, Silva FP, Li M, Yagyu R, Nakamura Y (2004) SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. Nat Cell Biol

Handwerger KE, Gall JG (2006) Subnuclear organelles: new insights into form and function. Trends Cell Biol 16:19-26

Harvey RP (1999) Seeking a regulatory roadmap for heart morphogenesis. Semin Cell Dev Biol 10:99-107

Hendzel MJ, Kruhlak MJ, Bazett-Jones DP (1998) Organization of highly acetylated chromatin around sites of heterogeneous nuclear RNA accumulation. Mol Biol Cell 9:2491-2507

Henikoff S (2005) Histone modifications: combinatorial complexity or cumulative simplicity? Proc Natl Acad Sci U S A 102:5308-5309

Henikoff S, Furuyama T, Ahmad K (2004) Histone variants, nucleosome assembly and epigenetic inheritance. Trends Genet 20:320-326

Henry KW, Berger SL (2002) Trans-tail histone modifications: wedge or bridge? Nat Struct Biol 9:565-566

Hess JL (2004) MLL: a histone methyltransferase disrupted in leukemia. Trends Mol Med 10:500-507

Horn PJ, Peterson CL (2002) Molecular biology. Chromatin higher order folding--wrapping up transcription. Science 297:1824-1827

Hoyt PR, Bartholomew C, Davis AJ, Yutzey K, Gamer LW, Potter SS, Ihle JN, Mucenski ML (1997) The Evi1 proto-oncogene is required at midgestation for neural, heart, and paraxial mesenchyme development. Mech Dev 65:55-70

Huang S (2002) Histone methyltransferases, diet nutrients and tumour suppressors. Nat Rev Cancer 2:469-476

Huang S, Shao G, Liu L (1998) The PR domain of the Rb-binding zinc finger protein RIZ1 is a protein binding interface and is related to the SET domain functioning in chromatin-mediated gene expression. J Biol Chem 273:15933-15939

Huyen Y, Zgheib O, Ditullio RA, Jr., Gorgoulis VG, Zacharatos P, Petty TJ, Sheston EA, Mellert HS, Stavridi ES, Halazonetis TD (2004) Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. Nature 432:406-411

Hwang I, Gottlieb PD (1995) Bop: a new T-cell-restricted gene located upstream of and opposite to mouse CD8b. Immunogenetics 42:353-361

Izutsu K, Kurokawa M, Imai Y, Ichikawa M, Asai T, Maki K, Mitani K, Hirai H (2002) The t(3;21) fusion product, AML1/Evi-1 blocks AML1-induced transactivation by recruiting CtBP. Oncogene 21:2695-2703

Jackson V (1988) Deposition of newly synthesized histones: hybrid nucleosomes are not tandemly arranged on daughter DNA strands. Biochemistry 27:2109-2120

Janicki SM, Spector DL (2003) Nuclear choreography: interpretations from living cells. Curr Opin Cell Biol 15:149-157

Jenuwein T, Allis CD (2001) Translating the histone code. Science 293:1074-1080

Jessell TM (2000) Neuronal specification in the spinal cord: inductive signals and transcriptional codes. Nat Rev Genet 1:20-29

Jones PA, Baylin SB (2002) The fundamental role of epigenetic events in cancer. Nat Rev Genet 3:415-428

Julien E, Herr W (2004) A switch in mitotic histone H4 lysine 20 methylation status is linked to M phase defects upon loss of HCF-1. Mol Cell 14:713-725

Kalantry S, Mills KC, Yee D, Otte AP, Panning B, Magnuson T (2006) The Polycomb group protein Eed protects the inactive X-chromosome from differentiation-induced reactivation. Nat Cell Biol 8:195-202

Karachentsev D, Sarma K, Reinberg D, Steward R (2005) PR-Set7-dependent methylation of histone H4 Lys 20 functions in repression of gene expression and is essential for mitosis. Genes Dev 19:431-435

Kavi HH, Fernandez HR, Xie W, Birchler JA (2005a) RNA silencing in Drosophila. FEBS Lett

Kavi HH, Xie W, Fernandez HR, Birchler JA (2005b) Global analysis of siRNA-mediated transcriptional gene silencing. Bioessays 27:1209-1212

Keats JJ, Maxwell CA, Taylor BJ, Hendzel MJ, Chesi M, Bergsagel PL, Larratt LM, Mant MJ, Reiman T, Belch AR, Pilarski LM (2005) Overexpression of transcripts originating from the MMSET locus characterizes all t(4;14)(p16;q32)-positive multiple myeloma patients. Blood 105:4060-4069

Kim KC, Geng L, Huang S (2003) Inactivation of a histone methyltransferase by mutations in human cancers. Cancer Res 63:7619-7623

Kim KC, Huang S (2003) Histone methyltransferases in tumor suppression. Cancer Biol Ther 2:491-499

Kizer KO, Phatnani HP, Shibata Y, Hall H, Greenleaf AL, Strahl BD (2005) A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. Mol Cell Biol 25:3305-3316

Kleefstra T, Smidt M, Banning MJ, Oudakker AR, Van Esch H, de Brouwer AP, Nillesen W, Sistermans EA, Hamel BC, de Bruijn D, Fryns JP, Yntema HG, Brunner HG, de Vries BB, van Bokhoven H (2005) Disruption of the gene Euchromatin Histone Methyl Transferase1 (Eu-HMTase1) is associated with the 9q34 subtelomeric deletion syndrome. J Med Genet 42:299-306

Kohlmaier A, Savarese F, Lachner M, Martens J, Jenuwein T, Wutz A (2004) A chromosomal memory triggered by Xist regulates histone methylation in X inactivation. PLoS Biol 2:E171

Kolomietz E, Marrano P, Yee K, Thai B, Braude I, Kolomietz A, Chun K, Minkin S, Kamel-Reid S, Minden M, Squire JA (2003) Quantitative PCR identifies a minimal deleted region of 120 kb extending from the Philadelphia chromosome ABL translocation breakpoint in chronic myeloid leukemia with poor outcome. Leukemia 17:1313-1323

Kourmouli K, Jeppessen P, Mahadevhaiah S, Burgoyne P, Wu R, Gilbert DM, Bongiorni S, Prantera G, Fanti L, Pimpinelli S, Shi W, Fundele R, Singh PB (2004) Heterochromatin and tri-methylated lysine 20 of histone H4 in animals. J Cell Sci 117:2491-2501

Kouskouti A, Scheer E, Staub A, Tora L, Talianidis I (2004) Gene-specific modulation of TAF10 function by SET9-mediated methylation. Mol Cell 14:175-182

Kouskouti A, Talianidis I (2005) Histone modifications defining active genes persist after transcriptional and mitotic inactivation. Embo J 24:347-357

Krogan NJ, Kim M, Tong A, Golshani A, Cagney G, Canadien V, Richards DP, Beattie BK, Emili A, Boone C, Shilatifard A, Buratowski S, Greenblatt J (2003) Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Mol Cell Biol 23:4207-4218

Kuenzle CC, Heizmann CW, Hubscher U, Hobi R, Winkler GC, Jaeger AW, Morgenegg G (1983) Chromatin changes accompanying neuronal differentiation. Cold Spring Harb Symp Quant Biol 48 Pt 2:493-499

Kurdistani SK, Tavazoie S, Grunstein M (2004) Mapping global histone acetylation patterns to gene expression. Cell 117:721-733

Kuwahara K, Barrientos T, Pipes GC, Li S, Olson EN (2005) Muscle-specific signaling mechanism that links actin dynamics to serum response factor. Mol Cell Biol 25:3173-3181

Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D (2002) Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev 16:2893-2905

Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature 410:116-120

Lachner M, O'Sullivan RJ, Jenuwein T (2003) An epigenetic road map for histone lysine methylation. J Cell Sci 116:2117-2124

Lamond AI, Earnshaw WC (1998) Structure and function in the nucleus. Science 280:547-553

Lamond AI, Spector DL (2003) Nuclear speckles: a model for nuclear organelles. Nat Rev Mol Cell Biol 4:605-612

Lande-Diner L, Cedar H (2005) Opinion: Silence of the genes - mechanisms of long-term repression. Nat Rev Genet

Lapunzina P (2005) Other tumors in Sotos syndrome. Am J Med Genet A 135:228

Larsson J, Chen JD, Rasheva V, Rasmuson-Lestander A, Pirrotta V (2001) Painting of fourth, a chromosome-specific protein in Drosophila. Proc Natl Acad Sci U S A 98:6273-6278

Lee MG, Wynder C, Cooch N, Shiekhattar R (2005) An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. Nature 437:432-435

Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, et al. (2006) Control of developmental regulators by Polycomb in human embryonic stem cells. Cell 125:301-313

Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AH (2003) Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol 13:1192-1200

Lin KT, Lu RM, Tarn WY (2004) The WW domain-containing proteins interact with the early spliceosome and participate in pre-mRNA splicing in vivo. Mol Cell Biol 24:9176-9185

Lin Q, Schwarz J, Bucana C, Olson EN (1997) Control of mouse cardiac morphogenesis and myogenesis by transcription factor MEF2C. Science 276:1404-1407

Lin W, Dent SY (2006) Functions of histone-modifying enzymes in development. Curr Opin Genet Dev 16:137-142

Liu H, Kim JM, Aoki F (2004) Regulation of histone H3 lysine 9 methylation in oocytes and early pre-implantation embryos. Development 131:2269-2280

Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389:251-260

Lynch J, Guo L, Gelebart P, Chilibeck K, Xu J, Molkentin JD, Agellon LB, Michalak M (2005) Calreticulin signals upstream of calcineurin and MEF2C in a critical Ca2+-dependent signaling cascade. J Cell Biol 170:37-47

Maison C, Bailly D, Peters AH, Quivy JP, Roche D, Taddei A, Lachner M, Jenuwein T, Almouzni G (2002) Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. Nat Genet 30:329-334

Malgeri U, Baldini L, Perfetti V, Fabris S, Vignarelli MC, Colombo G, Lotti V, Compasso S, Bogni S, Lombardi L, Maiolo AT, Neri A (2000) Detection of t(4;14)(p16.3;q32) chromosomal translocation in multiple myeloma by reverse transcription-polymerase chain reaction analysis of IGH-MMSET fusion transcripts. Cancer Res 60:4058-4061

Margueron R, Trojer P, Reinberg D (2005) The key to development: interpreting the histone code? Curr Opin Genet Dev 15:163-176

Marquardt T, Gruss P (2002) Generating neuronal diversity in the retina: one for nearly all. Trends Neurosci 25:32-38

Marquardt T, Pfaff SL (2001) Cracking the transcriptional code for cell specification in the neural tube. Cell 106:651-654

Marshall WF (2002) Order and disorder in the nucleus. Curr Biol 12:R185-192

Marshall WF, Straight A, Marko JF, Swedlow J, Dernburg A, Belmont A, Murray AW, Agard DA, Sedat JW (1997) Interphase chromosomes undergo constrained diffusional motion in living cells. Curr Biol 7:930-939

Martens JH, O'Sullivan RJ, Braunschweig U, Opravil S, Radolf M, Steinlein P, Jenuwein T (2005) The profile of repeat-associated histone lysine methylation states in the mouse epigenome. Embo J 24:800-812

McKinsey TA, Olson EN (2004) Cardiac histone acetylation--therapeutic opportunities abound. Trends Genet 20:206-213

McKinsey TA, Zhang CL, Olson EN (2001) Control of muscle development by dueling HATs and HDACs. Curr Opin Genet Dev 11:497-504

McKinsey TA, Zhang CL, Olson EN (2002) Signaling chromatin to make muscle. Curr Opin Cell Biol 14:763-772

McManus KJ, Biron VL, Heit R, Underhill DA, Hendzel MJ (2006) Dynamic changes in histone H3 lysine 9 methylations: identification of a mitosis-specific function for dynamic methylation in chromosome congression and segregation. J Biol Chem 281:8888-8897

Metzger E, Wissmann M, Yin N, Muller JM, Schneider R, Peters AH, Gunther T, Buettner R, Schule R (2005) LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature 437:436-439

Michalak M, Guo L, Robertson M, Lozak M, Opas M (2004) Calreticulin in the heart. Mol Cell Biochem 263:137-142

Mito Y, Henikoff JG, Henikoff S (2005) Genome-scale profiling of histone H3.3 replacement patterns. Nat Genet 37:1090-1097

Moat SJ, Doshi SN, Lang D, McDowell IF, Lewis MJ, Goodfellow J (2004a) Treatment of coronary heart disease with folic acid: is there a future? Am J Physiol Heart Circ Physiol 287:H1-7

Moat SJ, Lang D, McDowell IF, Clarke ZL, Madhavan AK, Lewis MJ, Goodfellow J (2004b) Folate, homocysteine, endothelial function and cardiovascular disease. J Nutr Biochem 15:64-79

Moen PT, Jr., Johnson CV, Byron M, Shopland LS, de la Serna IL, Imbalzano AN, Lawrence JB (2004) Repositioning of muscle-specific genes relative to the periphery of SC-35 domains during skeletal myogenesis. Mol Biol Cell 15:197-206

Monneret C (2005) Histone deacetylase inhibitors. Eur J Med Chem 40:1-13

Moran JL, Li Y, Hill AA, Mounts WM, Miller CP (2002) Gene expression changes during mouse skeletal myoblast differentiation revealed by transcriptional profiling. Physiol Genomics 10:103-111

Morillon A, Karabetsou N, Nair A, Mellor J (2005) Dynamic lysine methylation on histone H3 defines the regulatory phase of gene transcription. Mol Cell 18:723-734

Morishita K, Parganas E, Matsugi T, Ihle JN (1992) Expression of the Evi-1 zinc finger gene in 32Dc13 myeloid cells blocks granulocytic differentiation in response to granulocyte colony-stimulating factor. Mol Cell Biol 12:183-189

Nielsen SJ, Schneider R, Bauer UM, Bannister AJ, Morrison A, O'Carroll D, Firestein R, Cleary M, Jenuwein T, Herrera RE, Kouzarides T (2001) Rb targets histone H3 methylation and HP1 to promoters. Nature 412:561-565

Nishioka K, Rice JC, Sarma K, Erdjument-Bromage H, Werner J, Wang Y, Chuikov S, Valenzuela P, Tempst P, Steward R, Lis JT, Allis CD, Reinberg D (2002) PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin. Mol Cell 9:1201-1213

Nitta E, Izutsu K, Yamaguchi Y, Imai Y, Ogawa S, Chiba S, Kurokawa M, Hirai H (2005) Oligomerization of Evi-1 regulated by the PR domain contributes to recruitment of corepressor CtBP. Oncogene

Nolz JC, Gomez TS, Billadeau DD (2005) The Ezh2 methyltransferase complex: actin up in the cytosol. Trends Cell Biol 15:514-517

Nucifora G (1997) The EVI1 gene in myeloid leukemia. Leukemia 11:2022-2031

O'Carroll D, Erhardt S, Pagani M, Barton SC, Surani MA, Jenuwein T (2001) The polycomb-group gene Ezh2 is required for early mouse development. Mol Cell Biol 21:4330-4336

O'Carroll D, Scherthan H, Peters AH, Opravil S, Haynes AR, Laible G, Rea S, Schmid M, Lebersorger A, Jerratsch M, Sattler L, Mattei MG, Denny P, Brown SD, Schweizer D, Jenuwein T (2000) Isolation and characterization of Suv39h2, a second histone H3 methyltransferase gene that displays testis-specific expression. Mol Cell Biol 20:9423-9433

Ogawa H, Ishiguro K, Gaubatz S, Livingston DM, Nakatani Y (2002) A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells. Science 296:1132-1136

Okamoto I, Otte AP, Allis CD, Reinberg D, Heard E (2004) Epigenetic dynamics of imprinted X inactivation during early mouse development. Science 303:644-649

Pal-Bhadra M, Leibovitch BA, Gandhi SG, Rao M, Bhadra U, Birchler JA, Elgin SC (2004) Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. Science 303:669-672

Panagopoulos I, Isaksson M, Billstrom R, Strombeck B, Mitelman F, Johansson B (2003) Fusion of the NUP98 gene and the homeobox gene HOXC13 in acute myeloid leukemia with t(11;12)(p15;q13). Genes Chromosomes Cancer 36:107-112

Pardon JF, Wilkins MH (1972) A super-coil model for nucleohistone. J Mol Biol 68:115-124

Pavlath GK, Blau HM (1986) Expression of muscle genes in heterokaryons depends on gene dosage. J Cell Biol 102:124-130

Pereira SL, Grayling RA, Lurz R, Reeve JN (1997) Archaeal nucleosomes. Proc Natl Acad Sci U S A 94:12633-12637

Pereira SL, Reeve JN (1998) Histones and nucleosomes in Archaea and Eukarya: a comparative analysis. Extremophiles 2:141-148

Peters AH, Kubicek S, Mechtler K, O'Sullivan RJ, Derijck AA, Perez-Burgos L, Kohlmaier A, Opravil S, Tachibana M, Shinkai Y, Martens JH, Jenuwein T (2003) Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. Mol Cell 12:1577-1589

Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schofer C, Weipoltshammer K, Pagani M, Lachner M, Kohlmaier A, Opravil S, Doyle M, Sibilia M, Jenuwein T (2001) Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell 107:323-337

Phan D, Rasmussen TL, Nakagawa O, McAnally J, Gottlieb PD, Tucker PW, Richardson JA, Bassel-Duby R, Olson EN (2005) BOP, a regulator of right ventricular heart development, is a direct transcriptional target of MEF2C in the developing heart. Development 132:2669-2678

Plath K, Fang J, Mlynarczyk-Evans SK, Cao R, Worringer KA, Wang H, de la Cruz CC, Otte AP, Panning B, Zhang Y (2003) Role of histone H3 lysine 27 methylation in X inactivation. Science 300:131-135

Rayasam GV, Wendling O, Angrand PO, Mark M, Niederreither K, Song L, Lerouge T, Hager GL, Chambon P, Losson R (2003) NSD1 is essential for early post-implantation development and has a catalytically active SET domain. Embo J 22:3153-3163

Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, Schmid M, Opravil S, Mechtler K, Ponting CP, Allis CD, Jenuwein T (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature 406:593-599

Reid AG, Nacheva EP (2004) A potential role for PRDM12 in the pathogenesis of chronic myeloid leukaemia with derivative chromosome 9 deletion. Leukemia 18:178-180

Reinhart BJ, Bartel DP (2002) Small RNAs correspond to centromere heterochromatic repeats. Science 297:1831

Rice JC, Briggs SD, Ueberheide B, Barber CM, Shabanowitz J, Hunt DF, Shinkai Y, Allis CD (2003) Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. Mol Cell 12:1591-1598

Rice JC, Nishioka K, Sarma K, Steward R, Reinberg D, Allis CD (2002) Mitotic-specific methylation of histone H4 Lys 20 follows increased PR-Set7 expression and its localization to mitotic chromosomes. Genes Dev 16:2225-2230

Rosati R, La Starza R, Veronese A, Aventin A, Schwienbacher C, Vallespi T, Negrini M, Martelli MF, Mecucci C (2002) NUP98 is fused to the NSD3 gene in acute myeloid leukemia associated with t(8;11)(p11.2;p15). Blood 99:3857-3860

Roy S, Ng T (2004) Blimp-1 specifies neural crest and sensory neuron progenitors in the zebrafish embryo. Curr Biol 14:1772-1777

Sanders SL, Portoso M, Mata J, Bahler J, Allshire RC, Kouzarides T (2004) Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage. Cell 119:603-614

Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J, Kouzarides T (2002) Active genes are tri-methylated at K4 of histone H3. Nature 419:407-411

Santos-Rosa H, Schneider R, Bernstein BE, Karabetsou N, Morillon A, Weise C, Schreiber SL, Mellor J, Kouzarides T (2003) Methylation of histone H3 K4 mediates association of the Isw1p ATPase with chromatin. Mol Cell 12:1325-1332

Sarg B, Koutzamani E, Helliger W, Rundquist I, Lindner HH (2002) Postsynthetic trimethylation of histone H4 at lysine 20 in mammalian tissues is associated with aging. J Biol Chem 277:39195-39201

Sarma K, Reinberg D (2005) Histone variants meet their match. Nat Rev Mol Cell Biol 6:139-149

Sarmento OF, Digilio LC, Wang Y, Perlin J, Herr JC, Allis CD, Coonrod SA (2004) Dynamic alterations of specific histone modifications during early murine development. J Cell Sci 117:4449-4459

Scarano MI, Strazzullo M, Matarazzo MR, D'Esposito M (2005) DNA methylation 40 years later: Its role in human health and disease. J Cell Physiol 204:21-35

Schalch T, Duda S, Sargent DF, Richmond TJ (2005) X-ray structure of a tetranucleosome and its implications for the chromatin fibre. Nature 436:138-141

Schneider R, Bannister AJ, Kouzarides T (2002) Unsafe SETs: histone lysine methyltransferases and cancer. Trends Biochem Sci 27:396-402

Schneider R, Bannister AJ, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T (2004) Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. Nat Cell Biol 6:73-77

Schotta G, Ebert A, Krauss V, Fischer A, Hoffmann J, Rea S, Jenuwein T, Dorn R, Reuter G (2002) Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. Embo J 21:1121-1131

Schotta G, Ebert A, Reuter G (2003) SU(VAR)3-9 is a conserved key function in heterochromatic gene silencing. Genetica 117:149-158

Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G, Reinberg D, Jenuwein T (2004) A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes Dev 18:1251-1262

Schramke V, Sheedy DM, Denli AM, Bonila C, Ekwall K, Hannon GJ, Allshire RC (2005) RNA-interference-directed chromatin modification coupled to RNA polymerase II transcription. Nature 435:1275-1279

Schul W, van Driel R, de Jong L (1998) A subset of poly(A) polymerase is concentrated at sites of RNA synthesis and is associated with domains enriched in splicing factors and poly(A) RNA. Exp Cell Res 238:1-12

Seki Y, Hayashi K, Itoh K, Mizugaki M, Saitou M, Matsui Y (2005) Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. Dev Biol 278:440-458

Seligson DB, Horvath S, Shi T, Yu H, Tze S, Grunstein M, Kurdistani SK (2005) Global histone modification patterns predict risk of prostate cancer recurrence. Nature 435:1262-1266

Shen X, Collier JM, Hlaing M, Zhang L, Delshad EH, Bristow J, Bernstein HS (2003) Genome-wide examination of myoblast cell cycle withdrawal during differentiation. Dev Dyn 226:128-138

Shi YJ, Matson C, Lan F, Iwase S, Baba T, Shi Y (2005) Regulation of LSD1 histone demethylase activity by its associated factors. Mol Cell 19:857-864

Shilatifard A (2004) Transcriptional elongation control by RNA polymerase II: a new frontier. Biochim Biophys Acta 1677:79-86

Silva J, Mak W, Zvetkova I, Appanah R, Nesterova TB, Webster Z, Peters AH, Jenuwein T, Otte AP, Brockdorff N (2003) Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. Dev Cell 4:481-495

Sims RJ, 3rd, Weihe EK, Zhu L, O'Malley S, Harriss JV, Gottlieb PD (2002) m-Bop, a repressor protein essential for cardiogenesis, interacts with skNAC, a heart- and muscle-specific transcription factor. J Biol Chem 277:26524-26529

Solovei I, Schermelleh L, During K, Engelhardt A, Stein S, Cremer C, Cremer T (2004) Differences in centromere position of cycling and postmitotic cell types. Chromosoma 112:410-423

Spector DL (2001) Nuclear domains. J Cell Sci 114:2891-2893

Stancheva I (2005) Caught in conspiracy: cooperation between DNA methylation and histone H3K9 methylation in the establishment and maintenance of heterochromatin. Biochem Cell Biol 83:385-395

Stec I, Wright TJ, van Ommen GJ, de Boer PA, van Haeringen A, Moorman AF, Altherr MR, den Dunnen JT (1998) WHSC1, a 90 kb SET domain-containing gene, expressed in early development and homologous to a Drosophila dysmorphy gene maps in the Wolf-Hirschhorn syndrome critical region and is fused to IgH in t(4;14) multiple myeloma. Hum Mol Genet 7:1071-1082

Strahl BD, Allis CD (2000) The language of covalent histone modifications. Nature 403:41-45

Su IH, Dobenecker MW, Dickinson E, Oser M, Basavaraj A, Marqueron R, Viale A, Reinberg D, Wulfing C, Tarakhovsky A (2005) Polycomb group protein ezh2 controls actin polymerization and cell signaling. Cell 121:425-436

Szentirmay MN, Sawadogo M (2000) Spatial organization of RNA polymerase II transcription in the nucleus. Nucleic Acids Res 28:2019-2025

Tachibana M, Sugimoto K, Fukushima T, Shinkai Y (2001) Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. J Biol Chem 276:25309-25317

Tachibana M, Sugimoto K, Nozaki M, Ueda J, Ohta T, Ohki M, Fukuda M, Takeda N, Niida H, Kato H, Shinkai Y (2002) G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev 16:1779-1791

Taddei A, Hediger F, Neumann FR, Gasser SM (2004) The function of nuclear architecture: a genetic approach. Annu Rev Genet 38:305-345

Tagami H, Ray-Gallet D, Almouzni G, Nakatani Y (2004) Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. Cell 116:51-61

Tan X, Rotllant J, Li H, De Deyne P, Du SJ (2006) SmyD1, a histone methyltransferase, is required for myofibril organization and muscle contraction in zebrafish embryos. Proc Natl Acad Sci U S A 103:2713-2718

Tanabe Y, Jessell TM (1996) Diversity and pattern in the developing spinal cord. Science 274:1115-1123

Tatton-Brown K, Douglas J, Coleman K, Baujat G, Chandler K, Clarke A, Collins A, Davies S, Faravelli F, Firth H, Garrett C, Hughes H, Kerr B, Liebelt J, Reardon W, Schaefer GB, Splitt M, Temple IK, Waggoner D, Weaver DD, Wilson L, Cole T, Cormier-Daire V, Irrthum A, Rahman N (2005a) Multiple mechanisms are implicated in the generation of 5q35 microdeletions in Sotos syndrome. J Med Genet 42:307-313

Tatton-Brown K, Douglas J, Coleman K, Baujat G, Cole TR, Das S, Horn D, Hughes HE, Temple IK, Faravelli F, Waggoner D, Turkmen S, Cormier-Daire V, Irrthum A, Rahman N (2005b) Genotype-Phenotype Associations in Sotos Syndrome: An Analysis of 266 Individuals with NSD1 Aberrations. Am J Hum Genet 77:193-204

Tomczak KK, Marinescu VD, Ramoni MF, Sanoudou D, Montanaro F, Han M, Kunkel LM, Kohane IS, Beggs AH (2004) Expression profiling and identification of novel genes involved in myogenic differentiation. Faseb J 18:403-405

Tumbar T, Belmont AS (2001) Interphase movements of a DNA chromosome region modulated by VP16 transcriptional activator. Nat Cell Biol 3:134-139

Turner BM (2000) Histone acetylation and an epigenetic code. Bioessays 22:836-845

Turner BM (2002) Cellular memory and the histone code. Cell 111:285-291

Valk PJ, Verhaak RG, Beijen MA, Erpelinck CA, Barjesteh van Waalwijk van Doorn-Khosrovani S, Boer JM, Beverloo HB, Moorhouse MJ, van der Spek PJ, Lowenberg B, Delwel R (2004) Prognostically useful gene-expression profiles in acute myeloid leukemia. N Engl J Med 350:1617-1628

van Etten RW, de Koning EJ, Verhaar MC, Gaillard CA, Rabelink TJ (2002) Impaired NO-dependent vasodilation in patients with Type II (non-insulin-dependent) diabetes mellitus is restored by acute administration of folate. Diabetologia 45:1004-1010

van Leeuwen F, Gafken PR, Gottschling DE (2002) Dot1p modulates silencing in yeast by methylation of the nucleosome core. Cell 109:745-756

Vazquez J, Belmont AS, Sedat JW (2001) Multiple regimes of constrained chromosome motion are regulated in the interphase Drosophila nucleus. Curr Biol 11:1227-1239

Verdel A, Jia S, Gerber S, Sugiyama T, Gygi S, Grewal SI, Moazed D (2004) RNAi-mediated targeting of heterochromatin by the RITS complex. Science 303:672-676

Verdone L, Caserta M, Di Mauro E (2005) Role of histone acetylation in the control of gene expression. Biochem Cell Biol 83:344-353

Verhaar MC, Wever RM, Kastelein JJ, van Loon D, Milstien S, Koomans HA, Rabelink TJ (1999) Effects of oral folic acid supplementation on endothelial function in familial hypercholesterolemia. A randomized placebo-controlled trial. Circulation 100:335-338

Vermaak D, Ahmad K, Henikoff S (2003) Maintenance of chromatin states: an open-and-shut case. Curr Opin Cell Biol 15:266-274

Vidanes GM, Bonilla CY, Toczyski DP (2005) Complicated tails: histone modifications and the DNA damage response. Cell 121:973-976

Volpe TA, Kidner C, Hall IM, Teng G, Grewal SI, Martienssen RA (2002) Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science 297:1833-1837

Wakimoto BT (1998) Beyond the nucleosome: epigenetic aspects of position-effect variegation in Drosophila. Cell 93:321-324

Wang Y, Wysocka J, Sayegh J, Lee YH, Perlin JR, Leonelli L, Sonbuchner LS, McDonald CH, Cook RG, Dou Y, Roeder RG, Clarke S, Stallcup MR, Allis CD, Coonrod SA (2004) Human PAD4 regulates histone arginine methylation levels via demethylimination. Science 306:279-283

Wassenegger M (2005) The Role of the RNAi Machinery in Heterochromatin Formation. Cell 122:13-16

Waterborg JH (2002) Dynamics of histone acetylation in vivo. A function for acetylation turnover? Biochem Cell Biol 80:363-378

Weintraub H, Flint SJ, Leffak IM, Groudine M, Grainger RM (1978) The generation and propagation of variegated chromosome structures. Cold Spring Harb Symp Quant Biol 42 Pt 1:401-407

Weintraub H, Groudine M (1976) Chromosomal subunits in active genes have an altered conformation. Science 193:848-856

Weston AD, Rosen V, Chandraratna RA, Underhill TM (2000) Regulation of skeletal progenitor differentiation by the BMP and retinoid signaling pathways. J Cell Biol 148:679-690

Wiid IJ, Durrheim G, Bester AJ, van Helden PD (1988) Structural alterations in chromatin during myogenesis in the chicken. Mol Cell Biochem 79:57-62

Wilm TP, Solnica-Krezel L (2005) Essential roles of a zebrafish prdm1/blimp1 homolog in embryo patterning and organogenesis. Development 132:393-404

Witte RL, Godbout R (2002) Expression of spermidine/spermine N(1)-acetyltransferase in the Muller glial cells of the developing chick retina. Exp Eye Res 74:605-613

Wolk R (2003) Calcineurin, myocardial hypertrophy, and electrical remodeling. Cardiovasc Res 57:289-293

Wysocka J, Milne TA, Allis CD (2005) Taking LSD 1 to a new high. Cell 122:654-658

Xiao B, Jing C, Kelly G, Walker PA, Muskett FW, Frenkiel TA, Martin SR, Sarma K, Reinberg D, Gamblin SJ, Wilson JR (2005) Specificity and mechanism of the histone methyltransferase Pr-Set7. Genes Dev 19:1444-1454

Yamane K, Toumazou C, Tsukada YI, Erdjument-Bromage H, Tempst P, Wong J, Zhang Y (2006) JHDM2A, a JmjC-Containing H3K9 Demethylase, Facilitates Transcription Activation by Androgen Receptor. Cell

Yang XJ (2004) Lysine acetylation and the bromodomain: a new partnership for signaling. Bioessays 26:1076-1087

Yang XJ (2005) Multisite protein modification and intramolecular signaling. Oncogene 24:1653-1662

Yu BD, Hess JL, Horning SE, Brown GA, Korsmeyer SJ (1995) Altered Hox expression and segmental identity in Mll-mutant mice. Nature 378:505-508